Role of ATM in Suppressing Oncogenic Translocations and Mature B Cell Lymphomas

Role of ATM in Suppressing Oncogenic Translocations and Mature B cell Lymphomas

A dissertation presented

by

Suprawee Tepsuporn

to

The Division of Medical Sciences

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Immunology

Harvard University

Cambridge, Massachusetts

November 2013

Dissertation Advisor: Professor Frederick W. Alt Suprawee Tepsuporn

Role of ATM in Suppressing Oncogenic Translocations and Mature B cell Lymphomas

Abstract

The ATM protein senses DNA double-stranded breaks (DSBs) and facilitates proper repair. B and T lymphocytes of ATM-deficient patients have increased antigen receptor locus translocations associated with aberrant V(D)J recombination. Correspondingly, ATM-deficient humans are predisposed to both T and B cell malignancies. However, ATM-deficiency in mice only leads to T cell lymphomas, all of which harbor T cell receptor locus translocations resulting from aberrant V(D)J recombination.

The first goal of this study was to assess whether ATM-deficient B cell lymphomas occur in mice in which V(D)J recombination is targeted to the *c-myc* oncogene or in which B cell survival is increased via enforced expression of the anti-apoptotic Bcl2 protein. We found that, in the absence of ATM, either inserting the V(D)J substrate into *c-myc* or enhancing B cell survival led to the development of mature B cell lymphomas in a subset of mice. Moreover, combining both genetic alterations led to complete penetrance of mature B cell lymphomas in the ATM-deficient background.

The second goal was to elucidate the mechanisms that led to ATM-deficient B cell lymphoma development in our models. We found that nearly all tumors harbored recurrent chromosomal translocations that fused V(D)J-associated *IgH* locus DSBs to regions downstream

of *c-myc*, resulting in dicentric chromosomes and *c-myc* amplification by a breakage-fusion-bridge (BFB) mechanism. Then, we employed genome-wide translocation cloning to study the fate of introduced *c-myc* DSBs in activated wild-type and ATM-deficient B cells. In ATM-deficient, but not wild-type, B cells, a 30 megabase region downstream of *IgH* is a hotspot for *c-myc* translocations. We found that these translocation breakpoints originated from newly generated DSBs downstream of *IgH* in activated mature B cells and that many did not originate from the *IgH* class-switching process that introduces DSBs into *IgH* in mature B cells.

Based on our overall studies, we propose that ATM prevents conversion of unrepaired V(D)J-associated *IgH* chromosomal breaks in developing B cells into dicentric chromosomes that can persist through development. We further propose that, through a BFB mechanism, such dicentrics in mature B cells can give rise to new chromosome 12 DSBs that contribute to oncogenic translocations.

Table of Contents

ABSTRACT iii

TABLE OF CONTENTS v

DEDICATION viii

ACKNOWLEDGMENTS ix

CHAPTER 1: BACKGROUND 1

1.1. Study Overview 2

1.2. Antigen Receptor Gene Assembly in Developing Lymphocytes 11
1.2.1. Structure of Antibody Molecule and Immunoglobulin Loci Organization 11
1.2.2. Structure of T Cell Receptor Molecules and T Cell Receptor Loci Organization 13
1.2.3. Mechanisms of V(D)J Recombination 15
1.2.4. Developmental Regulation of V(D)J Recombination 16
1.2.5. Higher-Order Regulation of V(D)J Recombination 18

1.3. Secondary Antibody Diversification 24

1.4. The Role of DNA Repair Pathways in V(D)J Recombination and CSR 28
1.4.1 Classical Non-Homologous End Joining 28
1.4.2. Alternative End Joining 30
1.4.3. The ATM-Dependent DNA DSB Response 31
1.4.4. The p53-Dependent Cell Cycle Checkpoint 34

1.5. Chromosomal Aberrations and Cancers 36
1.5.1. Factors that Influence Oncogenic Chromosomal Translocation in Lymphocytes 37
1.5.2. Role of DSBs in Chromosomal Translocations 37
1.5.2.1. Role of RAG Activity in Translocations 38
1.5.2.2. AID-Initiated Translocations 41
1.5.2.3. Partnership of RAG and AID in Translocations 42
1.5.3. Role of Spatial Genome Organization in Translocation Frequency 43
1.5.4. Roles of Antigen Receptor Locus Enhancers in the Selection of Oncogenic Translocation 45

1.6. Role of DNA Repair in Suppressing Chromosomal Translocations 48
1.6.1. C-NHEJ/p53-Deficient Mouse Models 48

1.6.2. C-NHEJ Deficiency in Human 49
1.6.3. DNA Damage Response Deficiency in Mouse and Human 50

1.7. Cancer Development and ATM Deficiency 53
1.7.1. Clinical Features of Ataxia-telangiectasia 53
1.7.2. Cytogenetic Aberrations in AT Lymphocytes 53
1.7.3. Recurrent Translocations in ATM-Deficient Human Lymphoid Tumors 54
1.7.4. ATM Mutations in Sporadic Lymphomas 56
1.7.5. ATM Deficiency and Lymphoma Development in Mice 57
1.7.6. The Link Between ATM-Deficient Mouse Models and Cancer Development in AT Patients 60

CHAPTER 2: MATURE B CELL LYMPHOMA DEVELOPMENT IN ATM-DEFICIENT BACKGROUND 62

2.1. Summary 63

2.2. Introduction 63

2.3. Results 68
2.3.1. RAG Target Sequence in *c-myc* Lead to Mature B Cell Lymphoma in ATM-Deficient Mice 68
2.3.2. Bcl-2 Transgene Promotes Mature B Cell Lymphoma In ATM-Deficient Mice 71
2.3.3. ATM-Deficient Mature B Cell Lymphomas Arise From Naïve Mature B Cells Population 74
2.3.4. ATM-Deficient B Cell Lymphomas Harbor Complex Translocation Involving Chromosome 12 And 15 80
2.3.5. Translocation Breakpoints Downstream Of *c-myc* Lead To Gene Amplification Via Breakage-Fusion-Bridge Mechanism 86
2.3.6. Translocation In DA And DAB B Cell Lymphoma Do Not Involve RAG Target Sequence 93

2.4. Discussion 95
2.4.1. Mouse Models that Spontaneously Develop IgM+ B Cell Lymphomas in the Absence of ATM 95
2.4.2. Recurrent Oncogenic Translocations And Amplifications In ATM-Deficient B Cell Lymphomas. 96

CHAPTER 3: MECHANISMS OF CHROMOSOMAL TRANSLOCATION IN ATM-DEFICIENT PERIPHERAL B CELLS 100

3.1. Results 101
3.1.1. ATM Suppresses Genome-Wide Translocations In Stimulated B Cells 101
3.1.2. ATM Prevents The Formation Of Dicentrics 104

3.2. Discussion 109
3.2.1. RAG-Initiated DSBs May Persist Developmentally As Dicentrics And Give Rise To De Novo DSBs In Mature B Cells 109
3.2.2. Implications For ATM-Deficient Mouse B And T Cell Lymphomas.....112

CHAPTER 4: FUTURE DIRECTIONS.....115

CHAPTER 5: MATERIALS AND METHODS 121

BIBLIOGRAPHY 130

For Mom and Dad

With Love

ACKNOWLEDGMENT

First and foremost, I would like to thank Dr. Frederick Alt for supporting and mentoring me for the last several years. I am always impressed by his passion, his brilliance, his dedication to science, his drive to excel, his unmatched writing skills, and his cooking. It is a privilege to be trained by him. He taught me to think critically and to expect more from myself. I did not know on that first day I started in his laboratory in the summer of 2005 that I would grow and learn so much. Thank you so much for the opportunity to be a part of an exciting scientific discovery.

I would also like to thank Dr. Monica Gostissa for her advice and guidance. I greatly benefited from her scientific expertise and insights. Thank you for teaching me the many skills I learned in the past few years as well as for the countless number of stimulating discussions we had. It was wonderful working with her. I really admire her outlook on life. She is my inspiration.

Many thanks to Dr. Jiazhi Hu for his expertise and patience as we work on this project together. A heartfelt thank to current and past members of the Alt laboratory for their friendship, camaraderie, helpful discussions, assistance, and advice.

I would like to extend special thanks to my dissertation committee chair and members, Dr. Glenn Dranoff, Dr. Margaret Shipp, and Dr. A. Thomas Look, for their guidance, insights, and encouragement throughout my training. It is an honor to have them be a part of my scientific journey.

I would like to express my deepest gratitude to the Division of Medical Sciences and the Immunology Program, especially to Dr. David Cardozo and Leah Simons. They made the

unchartered Ph.D. path a much more enjoyable and pleasant one. Thank you for sharing your wisdom with me and for having my best interests at heart.

My special thanks and appreciation also go to Dr. Sharon Stranford for her incomparable mentorship, teaching, and research skills. I was very fortunate to be trained by her. The lessons I learned from working closely with Sharon were about independent thinking, teaching, perseverance, and how to become a scientist. She is a large part of why I pursued doctoral study in Immunology. It is my aspiration that I affect my students' lives as profoundly as Sharon affected mine. I would also like to thank Mount Holyoke College and the Department of Biological Sciences for an excellent education. Thank you for giving me the opportunity to learn, grow, and accomplish my dreams. I cherish my time at the College and could not have asked for a better place to call home.

I would also like to thank my friends. Thank you for being such great fans. I am very fortunate to have you as my friends. I particularly like to thank Dr. Jennifer Magee, Dr. Jana Mooster, and Dr. Alisha Weight for carefully reading this dissertation. To the super women in my life: Girija Goyal, Natalie Tuvana, Aleksandra Aljakna, I-Hsuan Wu, Dr. Rose Ndong, Dr. Rachael Bonawitz, and Dr. Maria Lim. I thank them for their friendship, wisdom, support, and encouragement through the tough times. They remind me what really matters in life, for which I am so grateful.

Next, I would like to thank my family. To my parents, I thank them for their unconditional love, sacrifice, support, and encouragement. Thank you for believing in me and for letting me leave home so young to pursue my dreams. I would not have achieved them without their support and understanding. To my sister, who make me laugh and brighten my

days. To my grandparents, whose patience and love carry me throughout the years. This dissertation is also dedicated to them. To my aunts, uncles, and cousins, for their love and support. Thank you for keeping me in touch with life at home, especially for sending me photos of my favorite foods. To the Horwitt family for loving me just like their own. My special thanks go to Rebecca for her superb scientific writing skills and for her editorial assistance in preparing this dissertation.

Lastly, I would like to thank my husband, Jed, for his extraordinary love and support. I am blessed to have him in my life. Thank you for taking care of me, for protecting me, and for being there for me. I would not be able to do this without him.

CHAPTER 1: BACKGROUND

1.1. Study Overview

B and T lymphocytes express specialized antigen receptors, respectively termed the B cell receptor (BCR) and the T cell receptor (TCR). The membrane-bound BCR and its secreted form, called antibodies, are comprised of a pair of identical immunoglobulin (Ig) heavy chains (IgH), encoded by the *IgH* locus, and a pair of identical Ig light chains (IgL), encoded by either the *Igκ* or *Igλ* light chain loci (Jackson et al., 2013). The N-terminal portion of the heavy and light chains contains the variable (V) region, and is involved in specific antigen binding, while the C-terminal portion is termed the constant (C) region (Cobb et al., 2006). The IgH C region (C_H) can be encoded by several different sets of C_H exons which determine Ig class and effector functions (Cobb et al., 2006). TCRs are made up of either αβ or γδ heterodimers encoded by TCR α/δ, β, and γ loci; the N-terminal region of TCR chains is also a variable region involved in antigen recognition (Cobb et al., 2006). Each individual B or T cell expresses a receptor with a unique specificity, with most foreign antigens being recognized by a subset B and/or T cells within the extremely large populations of unique B and T cell clones within a given individual (Vettermann and Schlissel, 2010).

Most of this great antibody and TCR diversity is not directly encoded in the genome, but is generated through a process in which the exons encoding Ig or TCR variable regions are assembled from component V, D, and J segments in developing B and T lymphocytes by a process termed V(D)J recombination (Jackson et al., 2013). The V(D)J recombination process is initiated in progenitor B cells in the bone marrow (BM) or progenitor T cells in the thymus by a lymphocyte-specific endonuclease comprised of the recombination activating gene 1 and 2 proteins, which form a complex referred to as RAG (Nishana and Raghavan, 2012). RAG initiates

the cleavage phase of V(D)J recombination by introducing DNA double-strand breaks (DSBs) at the borders of a pair of V, D, or J segments (termed coding segments) and short conserved flanking sequences referred to as recombination signal sequences (RSSs) (Schatz et al., 2008). These RAG-initiated DSBs at the two participating coding segments and their flanking RSSs are joined exclusively by the classical non-homologous DNA end-joining (C-NHEJ) pathways, resulting in V(D)J coding joins and RSS joins, respectively (Boboila et al., 2012). V(D)J recombination is initiated by RAG and completed by C-NHEJ strictly in the G1 phase of the cell cycle (Helmink and Sleckman, 2012).

During B cell development, complete V(D)J exons are assembled upstream of the Cμ constant region exons, the first set of C_H exons expressed during development. Transcription is initiated upstream of the V(D)J and terminated downstream of the Cμ exons to produce a transcript that after splicing generates a messenger RNA encoding the μ heavy chain protein (Perlot and Alt, 2008). Subsequent assembly and expression of a functional IgL chain gene, which associates with the μ heavy chain, leads to the formation of IgM, which is expressed on the surface of the resulting IgM^+ B cell (Vettermann and Schlissel, 2010). Surface IgM-expressing B cells migrate from the BM to peripheral lymphoid organs, such as the spleen or lymph nodes, where upon stimulation with a cognate antigen or through certain of types of activation, they can be stimulated to undergo two additional Ig diversification processes, namely IgH class switch recombination (CSR) and IgH and IgL somatic hypermutation (SHM) (Di Noia and Neuberger, 2007). CSR changes antibody effector functions by replacing the Cμ exons with a different set of C_H exons from sets of C_H exons that lie 100-200kb downstream in the *IgH* locus (Keim et al., 2013). SHM, which occurs in specialized lymphoid structures termed germinal

centers (GCs), introduces point mutations into IgH and IgL variable region exons, generating further BCR variable region diversity in a given B cell clone and allowing for selection of B cells that express BCRs with higher antigen-binding affinity, a process termed affinity maturation (Di Noia and Neuberger, 2007).

CSR and SHM are initiated by activation-induced cytidine deaminase (AID), a protein that deaminates cytidine residues on single-stranded DNA (ssDNA) (Chaudhuri et al., 2007). This deamination process initiates a cascade of reactions that involve activities of normal base excision (BER) and mis-match repair (MMR) pathways that, instead of normal repair, are co-opted to introduce mutations or DSBs (Di Noia and Neuberger, 2007). For CSR, the AID-initiated DSBs occur in large repetitive sequences, known as switch (S) regions, which precede each set of C_H exons. The DSBs in the donor Sμ region are joined to the DSBs in one of the acceptor S regions preceding downstream set of C_H exons (Chaudhuri et al., 2004). To complete CSR, these AID-initiated S region DSBs also are usually joined by C-NHEJ (Boboila et al., 2012). However, in the absence of C-NHEJ factors, S region DSBs can still be fused to yield CSR by alternative end-joining (A-EJ) pathways, although to a lessor extent than with normal C-NHEJ (Boboila et al., 2012). During SHM, AID-initiated lesions in variable region exons are predominantly processed into mutations as opposed to DSBs, with this mutation activity also involving co-opted BER and MMR pathway activity (Di Noia and Neuberger, 2007). CSR also is initiated during the G1 cell cycle phase and completed predominantly by C-NHEJ of two region DSBs (Nussenzweig and Nussenzweig, 2010)

Chromosomal translocations appear to result predominantly from joining two separate DSBs together. Translocations can be inter-chromosomal if the join DSBs that occur on two

different chromosomes. Depending on how the DSBs are joined, inter-chromosomal translocations can result in the joining of the telomeric end of one chromosome to the centromeric end of another or they can result in dicentric chromosomes and/or acentric chromosome fragments (Bunting and Nussenzweig, 2013). Translocations can also occur intra-chromosomally if they join two separate DSBs on the same chromosome; such intrachromosomal translocations might result in deletions or inversions depending which ends of the DSBs are joined (Zhang et al., 2010). Joining of a single DSBs that has been resected can also result in deletions. The high frequency of DSBs generated during B and T cell development and activation pose a great threat for the generation of translocations and other chromosomal aberrations (Gostissa et al., 2011). Indeed, spontaneous B or T cell lymphomas in humans and certain mouse models often harbor recurrent translocations that fuse Ig or TCR loci to oncogenes and/or lead to deletion of tumor suppressor genes (Seifert et al., 2013). A classic example is the, T(8;14) translocation in Burkitt's B cell lymphoma (BL) that joins the *c-MYC* oncogene on chromosome 8 to the *IGH* locus on chromosome 14, thereby bringing *c-MYC* under the control of the 3' regulatory elements of *IGH* leading to the deregulation of *c-MYC* (Robbiani and Nussenzweig, 2013; Gostissa et al., 2011).

C-NHEJ maintains genomic integrity by promoting re-joining of DSBs and, thereby, suppressing chromosome breaks and the joining separate DSBs to form chromosomal translocations (Gostissa et al., 2011). C-NHEJ works throughout the cell cycle but is most important in G1 when homologous recombination, the other major mammalian DSB repair pathway, is not available (Boboila et al., 2012). Ku70, Ku80, XRCC4, and Ligase 4 are "core" C-NHEJ factors and are absolutely required for V(D)J recombination (Boboila et al., 2012). Ku70-

or Ku80-deficient mice exhibit growth defects, increased sensitivity to DNA damaging agents, chromosomal instability, and severe combined immunodeficiency (SCID) due to an inability to carry out end-joining to complete V(D)J recombination (Gostissa et al., 2011). XRCC4 or Lig4 deficiency results in late embryonic lethality due to neuronal lethality; but developing embryos exhibit similar defects to those associated with Ku-deficiency including a block in B and T cell development due to a V(D)J recombination defect (Gao et al., 2000; Frank et al., 1998). Most C-NHEJ deficient mice do not recurrently develop lymphoid or other cancers, at least in part due to cells containing unrepaired G1 DSBs being eliminated at the G1/S phase transition via cell cycle checkpoints (Puebla-Osorio and Zhu, 2008). In this regard, combined deficiency in mice for a C-NHEJ component and p53, a tumor suppressor which activates the G1 cell cycle checkpoint in response to unrepaired DSBs, leads to recurrent development of pro-B cell lymphomas in mice (Gostissa et al., 2011). A hallmark of pro-B cell lymphomas in C-NHEJ/p53-double deficient background is the presence of complex translocations with gene amplification, referred to as complicons. The complicon derives from the fusion of RAG-initiated DSBs in the *IgH* locus on chromosome 12 to a large region downstream of the *c-myc* gene on chromosome 15 generating dicentric chromosomes that ultimately lead to oncogenic amplification of the *c-myc* genes via breakage-fusion-bridge (BFB) mechanism (Gostissa et al., 2011; Zhu et al., 2002). Mechanistically, these complicons are thought to be result from unrepaired RAG-initiated DSBs due to C-NHEJ deficiency persisting through the G1 checkpoint in the absence of p53 and then being replicated to generate dicentrics (Difilippantonio et al., 2002; Zhu et al., 2002).

The Ataxia Telangiectasia-mutated (ATM) DNA damage response (DDR) pathway is activated in response to DSBs in G1, including those initiated by RAG and AID, to further

promote efficient C-NHEJ DSB repair and also to effect protective G1 checkpoints. In response to DSBs, the ATM kinase is activated and then, in turn, activates both downstream components involved in promoting repair of DSBs by C-NHEJ and in mediating checkpoints (Bednarski and Sleckman, 2012). In this context, in developing lymphocytes, the ATM DSB response stabilizes RAG-initiated V(D)J recombination DSBs and, more generally, appears to tether DSBs to prevent premature separation, thereby, contributing is required for their proper repair by C-NHEJ (Huang et al., 2007; Vacchio et al., 2007; Bredemeyer et al., 2006, Matei et al., 2006). In response to DSBs in G1, ATM also activates p53 to enforce the G1 cell cycle checkpoint to arrest cells to allow for proper DSB repair and/or to eliminate cells with persistent unrepaired DSBs before progression into S phase (Helmink and Sleckman, 2012). As mentioned above, the p53-dependent G1 checkpoint prevents cells from replicating unrepaired DSBs which could lead to chromosomal aberrations and further genome instability (Gostissa et al., 2011).

ATM-deficiency in humans leads to Ataxia Telangiectasia (AT), a condition characterized by neurodegeneration, immunodeficiency, increased sensitivity to ionizing radiation, and increased predisposition to T and B cell malignancy (Perlman et al., 2012). ATM inactivation in mice recapitulates several aspects of AT, including the predisposition to thymic lymphoma (Ambrose and Gatti, 2013). Notably, murine ATM-deficient thymic lymphomas all harbor recurrent chromosomal translocations that all involve RAG-initiated DSBs in the *Tcrd* locus on chr 14 (Zha et al., 2010). These recurrent *Tcrd* translocations include the generation of complicons from chr 14 dicentrics that are initiated by RAG-generated DSBs in the *Tcrd* locus and which are accompanied by joining of *Tcrd* chr 14 DSBs to DSBs on other chromosomes, most notably chr 12 downstream of the IgH locus (Zha et al., 2010). This picture of

chromosomal aberrations in ATM-deficient thymic lymphomas is strikingly reminiscent of the chromosomal aberrations found in the C-NHEJ/p53-deficient pro-B cell lymphomas (see above). Mechanistically, such similarities may be explained by the fact that ATM deficiency provides both the V(D)J joining defect and the G1-checkpoint defect (Helmink and Sleckman, 2012). In this regard, it is also notable that this "double whammy" of ATM-deficiency allows V(D)J recombination-associated DSBs within the *IgH* locus in developing pro-B cells to somehow persist through development into mature B cells where they can contribute to translocations (Callen et al., 2007).

Despite the genomic instability and predisposition to T cell lymphoma, ATM-deficient mice do not develop B lineage lymphomas as are often observed in human AT patients. The initial goal of the studies described in this thesis was to create a mouse model of mature B cell lymphoma in the ATM-deficient background. We hypothesized that the frequency of DSBs within oncogenes such as *c-myc* and/or the survival of ATM-deficient B cells may be rate-limiting factors that prevent the accumulation of oncogenic translocations and the development of B cell lymphoma. To test the role of DSBs, we asked whether the targeting of RAG-initiated DSBs to *c-myc* locus in developing mouse B cells could lead to increased predisposition to B cell malignancies with *c-myc* translocations. To test the role of B cell survival, we asked whether enforced expression of the bcl-2 anti-apoptotic factor in ATM-deficient B cells *in vivo* leads to B cell lymphoma development. We found that, in the absence of ATM, either inserting RAG target sequences into the *c-myc* locus or enhancing B cell survival via enforced Bcl2 expression led to the development of peripheral, predominantly IgM^+, B cell lymphomas in a significant fraction of the mice. Strikingly, combining both RAG targets in *c-myc*

and deregulated Bcl2 expression led to complete penetrance of such peripheral B cell lymphomas.

The second goal of our studies, was to determine whether the ATM-deficient mature B cell lymphomas that arose in our model were associated with recurrent genomic aberrations and, if so, to determine the underlying mechanisms. In this regard, we found that nearly all of the ATM-deficient mature B cell lymphomas that arose in each of three ATM deficient models harbored translocations that appeared to fuse RAG-initiated DSBs during attempted V(D)J joining on chromosome 12 within the *IgH* locus to DSBs on chromosome 15 that occurred within a region 5kb to several hundred kb downstream of *c-myc*. Such joins, as observed in C-NHEJ and p53 deficient pro-B lymphomas, result in the formation of dicentric chromosomes and *c-myc* amplification by the breakage-fusion-bridge mechanism. To elucidate mechanisms of how RAG-generated DSBs might lead to dicentric chromosome formation and oncogenic translocations in peripheral (mature) B cells, we employed high throughput genome-wide translocation sequencing technology (HTGTS) to study the fate of DSBs introduced into *c-myc* gene in wild-type and ATM-deficient splenic B cells activated for CSR. Mapping of genome-wide translocation junctions revealed that in both WT and ATM-deficient B cells DSBs in *c-myc* frequently translocate to AID-initiated DSBs in donor and acceptor IgH S regions as previously observed (Chiarle et al., 2011). However, in ATM-deficient B cells but not wild type B cells, *c-myc* translocation junctions also occurred at high levels in a 30 megabase region downstream of the *IgH* locus. Further analysis demonstrated that many translocations downstream of the *IgH* locus in ATM-deficient cells were generated *de novo* and in an AID-independent manner.

Based on the results of our overall studies and those of others (Bunting and Nussenzweig, 2013), we propose that the *IgH* DSBs in ATM-deficient peripheral B cells resulted from BFB cycles involving dicentrics of chromosomes 12 and that such dicentrics are formed both during aberrant repair of AID-initiated and RAG-initiated DSBs. We further propose that, in the absence of normal G1 checkpoints due to ATM-deficiency, dicentrics arising from RAG-initiated DSBs generated in pro-B cells are propagated through development to mature B cells where they can form new DSBs via BFB mechanism that can contribute to translocations in peripheral B cells. Together, findings from our tumor models and our HTGTS studies show a novel role for ATM in suppressing unrepaired RAG-initiated DSBs from propagating during development and serving as substrates for the formation of dicentric chromosomes and further DSBs in peripheral B cells that can contribute to oncogenic translocation. These findings may be relevant to human peripheral B cell lymphomas with apparent V(D)J recombination associated translocations.

1.2. Antigen Receptor Gene Assembly In Developing Lymphocytes

1.2.1. Structure Of Antibody Molecule And Immunoglobulin Loci Organization

The IgH and IgL chains are comprised of variable and constant regions (Cobb et al., 2006). These regions are composed of three-dimensionally folded segments called domains. The Ig heavy chain consists of one variable (V_H) and three constant (C_H) domains. The Ig light chain consists of one variable (V_L) and one constant (C_L) domain (Jackson et al., 2013). Each light chain is bound to the heavy chain by a disulfide bond and by a combination of noncovalent interactions such as salt bridges, hydrogen bonds, and hydrophobic bonds (Nishana and Raghavan, 2012). The two heavy chains are held together by similar noncovalent and disulfide bonds, resulting in the four-chain immunoglobulin structure (Figure 1).

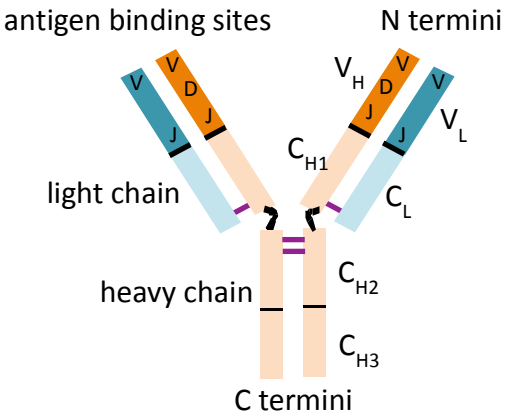


Figure 1. Schematic representation of an immunoglobulin. The IgH chains are depicted in orange with the variable region in dark orange. The IgL chains are depicted in blue with variable region in dark blue. Each domain is depicted as rectangles. Assembled V(D)J exons are indicated. Disulfide bonds are depicted as purple lines.

The IgH chain is encoded by variable (V), diversity (D), and joining (J) gene segments. In mouse, the *IgH* locus spans approximately 3 Mb near the telomeric portion of chromosome 12 with several hundred VH gene segments, 10-13 DH gene segments, and 4 JH gene segments,

followed by 8 C_H genes (Cμ, Cδ, Cγ3, Cγ1, Cγ2b, Cγ2a, Cε, and Cα, from 5' to the 3' of the locus) (Shih and Krangel, 2013). VH segments that lie closer to the D region are termed "proximal," whereas VH segments that lie further upstream are termed "distal" (Jung et al., 2006). A strong enhancer element termed the intronic enhancer (iEμ) lies between JH 4 and Cμ, the first CH gene, and function to promote V(D)J recombination (Gostissa et al., 2011). A series of enhancers termed the IgH 3' regulatory region (IgH 3'RR) lies downstream of Cα, the most 3' CH gene, and regulates CSR in mature B cells by modulating *IgH* locus conformation and looping (see below) (Gostissa et al., 2011). Transcriptional promoters are located upstream of each VH segment, upstream of DH segments, and in iEμ (Cobb et al., 2006).

The murine *Igκ* locus on chromosome 6 comprises more than 100 Vκ and 4 functional Jκ gene segments, with a single Cκ gene. The κ intronic enhancer (iEκ) lies between the Jκ cluster and the Cκ gene and the 3' κ enhancer (3'Eκ) lies downstream of the Cκ gene (Shih and Krangel, 2013). The murine *Igλ* locus is arranged in two duplicated units and spans approximately 200 kb on chromosome 16 (Cobb et al., 2006). The λ locus contains a total of three functional Vλ and Jλ gene segments, and three Cλ genes with each Jλ segment directly upstream of the respective Cλ gene, and a λ enhancer (Eλ) located downstream of each λ locus unit (Figure 2).

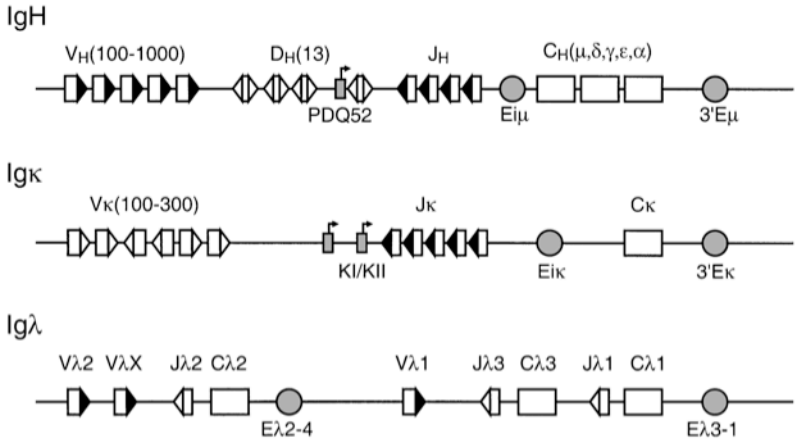


Figure 2. Schematic representation of the murine B cell receptor loci. The V, D, and J gene segments are depicted as rectangles. The estimated number of antigen receptor gene segments for the V_H and Vκ loci is indicated above each locus. The 12-RSSs are depicted as open triangles. The 23-RSSs are depicted as solid triangles. Enhancers and promoters are depicted as gray circles and rectangles, respectively. Constant regions are depicted as single rectangles without showing individual exons. Distances between the various elements are not drawn to scale. From Hesslein and Schatz, 2001.

1.2.2. Structure Of T Cell Receptor Molecules And T Cell Receptor Loci Organization

The TCR shares common characteristics with the immunoglobulin molecule. It is comprised of two different polypepetide chains, αβ or γδ. Each chain consists of an N-terminal variable (V) domain, followed by a constant (C) domain, and then a membrane-anchoring domain (Zoete et al., 2013). The V domains from the two chains organize into one antigen-binding site. Unlike B cells, after T cells are stimulated with antigen, there is no further mutation in the antigen-binding site or switching of constant-region class as occurs for immunoglobulins.

TCRs are encoded by β, δ, α, and γ gene families, organized in three genomic loci: α/δ, β, and γ (Cobb et al., 2006). Similar to IgH, the TCRβ and TCRδ variable region exons are assembled from V, D, and J gene segments. On the other hand, the TCRα and TCRγ variable

region exons are assembled from only V and J gene segments (Cobb et al., 2006). The TCRβ locus spans approximately 700 kb region on mouse chromosome 6 and contains two Cβ genes, each associated with one Dβ and six Jβ gene segments, and 34 Vβ gene segments located upstream of the DJβ clusters (Cobb et al., 2006).

The TCRα and δ loci are embedded in the same genomic region on chromosome 14. In these loci, there are 12 Vδ, two Dδ and two Jδ gene segments, followed by one Cδ gene, and 85 Vα and at least 60 Jα gene segments upstream of a single Cα gene (Cobb et al., 2006). The δ locus is located between the α V and J regions so that it is deleted in cells undergoing TCRα V-to-J recombination (Krangel, 2009).

The TCRγ locus, spans approximately 200 kb on mouse chromosome 13 and consists of seven Vγ and three functional Jγ and Cγ genes (Cobb et al., 2006) (Figure 3). The organization of human *Ig* and *TCR* loci is similar to their mouse counterparts (Hesslein and Schatz, 2001).

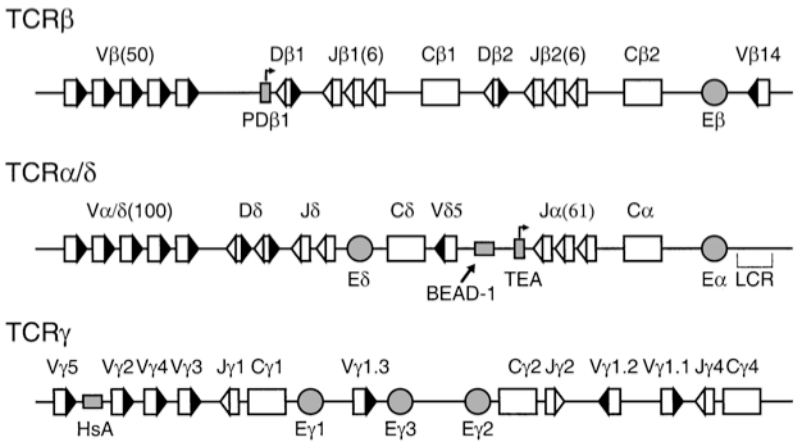


Figure 3. Schematic representation of the murine T cell receptor loci. The V, D, and J gene segments are depicted as rectangles. The estimated number of antigen receptor gene segments for each locus is indicated above each locus. The 12-RSSs are depicted as open triangles. The 23-RSSs are depicted as solid triangles. Enhancers and promoters are depicted as gray circles and rectangles, respectively. Constant regions are depicted as single rectangles. Distances between the various elements are not drawn to scale. From Hesslein and Schatz, 2001.

1.2.3. Mechanisms Of V(D)J Recombination

V(D)J recombination is initiated when developing lymphocytes express the RAG endonuclease (Nishana and Raghavan, 2012). RAG initiates V(D)J recombination by introducing DSBs between V, D, and J segments and their flanking RSSs, resulting in a pair of hairpin-sealed coding ends and a pair of blunt signal ends (Schatz and Swanson, 2011). RAG also functions to hold these DSBs in a postcleavage synaptic complex until they can be joined by C-NHEJ (Helmink and Sleckman, 2012). RSSs contain highly conserved palindromic 7-bp sequence (heptamer) and 9-bp sequence (nonamer) motifs, separated by a relatively nonconserved spacer of either 12 or 23 bp, referred to as "12RSSs" and "23RSSs", respectively (Nishana and Raghavan, 2012).

RAG function requires simultaneous binding to two RSSs, one with a 12RSS and the other with a 23RSS, a restriction referred to as the 12/23 rule (Brady et al., 2010). In IgH, for example, the V_H segments are flanked by 23RSSs, 12RSSs flank both sides of the D_H segments, and 23RSSs flank the JH segments. The 12/13 rule contributes to the ordered rearrangement of receptor loci in developing B cells such that the assembly of the D- JH segments occurs before they recombine to V_H segments, thereby preventing direct V- JH recombination. In contrast, the configuration of the TCRβ locus is such that 23RSSs flank Vβs, 12RSSs flank Jβs, and 12 and 23RSSs flank Dβs at either side. However, V(D)J recombination still progresses first with the assembly of the D-Jβ segment then the V-DJβ segment. This restriction, which is still mediated by RSS sequence, is referred to as "beyond 12/23 rule" (Brady et al., 2010). In this regards, several studies have shown that the beyond 12/23 rule is enforced by specific RSSs (Tillman et al., 2004; Jung et al., 2003; Wu et al., 2003; Bassing et al., 2000). Restricting RAG expression to developing lymphocytes and targeting RAG-dependent cutting to specific RSS sequences are

important mechanisms to prevent generation of off-target DSBs, which could potentially have a major impact on genomic stability (see below).

1.2.4. Developmental Regulation Of V(D)J Recombination

To ensure proper cleavage and joining of the coding segments, the recombination and assembly of functional V(D)J exons is a tightly regulated phenomenon. V(D)J recombination in developing B and T cells takes place in a lineage-specific manner: despite being subjected to the same V(D)J recombination machinery, functional rearrangement of *Ig* loci only takes place in B cells and functional rearrangement of *Tcr* loci only in T cells (Cobb et al., 2006). Among cells of the same lineage, V(D)J recombination is further regulated in a developmental stage-specific manner by regulated gene accessibility, which confers ordered rearrangement of antigen receptor genes such that the *IgH* locus undergoes rearrangement before the *IgL* loci in developing B cells and the *Tcrb* locus undergoes rearrangements before the *Tcra* locus in developing αβ T cells (Cobb et al., 2006).

Early B lineage cells, termed pro-B cells, activate the *IgH* locus and first recombine D_H and JH gene segments, followed by V_H to DJH rearrangement to generate a complete VH(D)JH exon (Vettermann and Schlissel, 2010). DH-to-JH recombination occurs on both *IgH* alleles, whereas the subsequent VH to DJH rearrangement occurs only on one allele, a process called allelic exclusion, which prevents the assembly of multiple antigen receptors in a single cell (Luning Prak et al., 2011). Given that the nature of coding joins is inherently imprecise, only one in three rearrangements will give rise to an open reading frame capable of expressing a functional protein. Therefore, B cells which first assemble a non-productive or non-pairing (see

below) VH(D)JH exon can undergo rearrangement of the second *IgH* allele (Vettermann and Schlissel, 2010) (Figure 4).

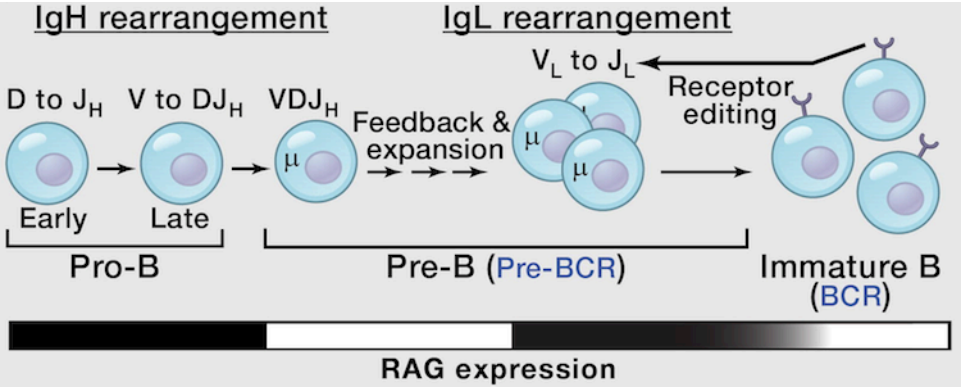


Figure 4. Schematic representation of mouse B cell development. The stage-specific rearrangement of *IgH* and *IgL* loci are indicated. The expression of RAG is depicted and darker shading indicates a higher relative level of expression. From Alt et al., 2013.

Upon productive rearrangement of the V(D)J exon, a μ heavy chain (μHC) encoded by V(D)J and Cμ is produced. The μ heavy chain is paired with the surrogate light chain (SLC), which is comprised of the invariant polypeptides VpreB and λ5 (Vettermann and Schlissel, 2010). Successful assembly of μHC to SLC produces a pre-B-cell receptor (pre-BCR) that signals differentiation of pro-B cells to the pre-B cell stage, where the *IgL* genes are assembled. *Igκ* rearrangement always precedes *Igλ* rearrangement (Luning Prak et al., 2011). Productive rearrangement of Vκ and Jκ gene segments allows for the expression of Igκ light chain (κLC) that can be paired with μHC to form a functional BCR. An inability to pair κLC with the preexisting μHC or a pairing that results in an autoreactive BCR leads to secondary rearrangements of the *IgL* loci via a process called receptor editing (Luning Prak et al., 2011). During receptor editing, rearranged VκJκ exons can be replaced by rearrangement of upstream Vκ to downstream Jκ gene segments (Luning Prak et al., 2011). Alternatively, a rearrangement

that deletes the *Igκ* locus allows the cells to undergo rearrangement in the *Igλ* locus. Expression of a functional BCR on the cell surface allows pre-B cells to develop into IgM^+ immature B cells, which migrate to the periphery.

Developing T cells follows similar regulation as they undergo rearrangement and expression of various *Tcr* genes in the thymus. Rearrangement of TCR α/δ, β, and γ are carried out at two distinct stages of thymocyte development. Thymocytes commit to the αβ or γδ lineages upon the outcome of these TCR gene recombination events. In the case of αβ T cells, the *Tcrb* locus rearranges first in the $CD4^-CD8^-$ double-negative (DN) 2 and 3 stages of thymocyte development (Krangel, 2009). Developmental stage-specific recombination of *Tcrb* dictates that D_β-to-J_β rearrangement precedes V_β-to-DJ_β recombination and the expression of a productive rearrangement of the *Tcrb* gene allows for differentiation into $CD4^+CD8^+$ double-positive (DP) cells, in which *Tcrα* recombination occurs (Spicuglia et al., 2010). Afterwards, TCRαβ-expressing cells can be further selected into mature $CD4^+$ or $CD8^+$ single-positive (SP) T-cells (Figure 5). Similarly to *IgH*, the *TCRβ* locus is allelically excluded, so that only one allele is functionally rearranged and expressed in mature αβ T cells (Luning Prak et al., 2011; Krangel, 2009).

1.2.5. Higher-Order Regulation of V(D)J Recombination

The correct execution of V(D)J recombination is tightly controlled by several higher-order mechanisms that include transcription, chromatin modifications, and chromosome conformational changes, which ultimately act by modulating locus accessibility. A first clue to this regulatory mechanism begins with the discovery of germline transcripts. It is observed that

transcription of germline gene segments coincided with B cell developmental stages where these gene segments are poised for recombination (Yancopoulos and Alt, 1985). Since these initial observations, the correlation between germline transcription and V(D)J recombination has been extended to all Ig and TCR gene segments (Abbey and O' Neill, 2008).

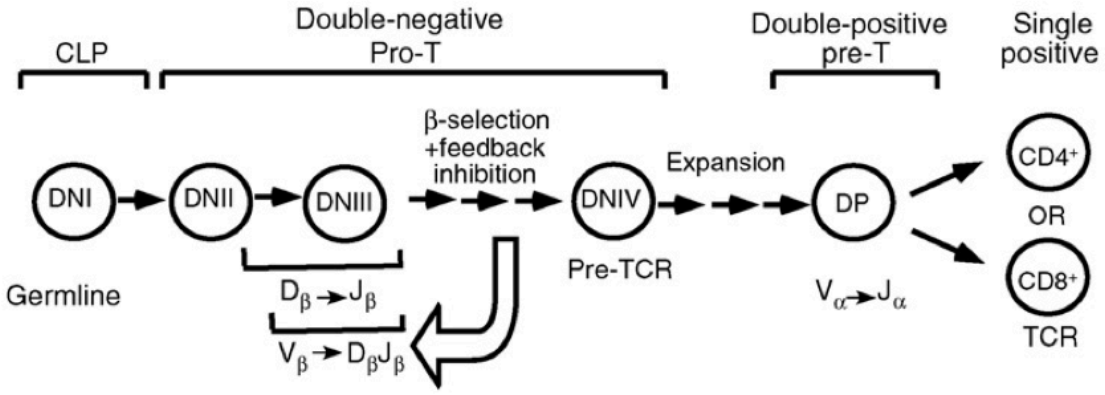


Figure 5. Schematic representation of mouse α/β T cell development. The developmental stage-specific rearrangements of TCR genes and surface expression of the pre-TCR are indicated. From Cobb et al., 2006.

Genic and intergenic germline antisense transcription occurs extensively in the V_H region (Matheson and Corcoran, 2012). This transcription occurs on both alleles and was detected in a large proportion of pro-B cells prior to V_H-to-DJ_H rearrangement, suggesting that this antisense transcription remodels the V_H region to promote accessibility to the VH gene segments (Bolland et al., 2004). Lastly, in pre-B and-mature B cells, sense transcription of assembled $V(D)J_H$ exons is detected from the V_H promoter but neither sense nor antisense transcription occurs in the distal V_H segments (Giallourakis et al., 2004) (Figure 6). Taken together, transcriptional status of Ig genes may play an important role in opening up the locus to recombination machinery.

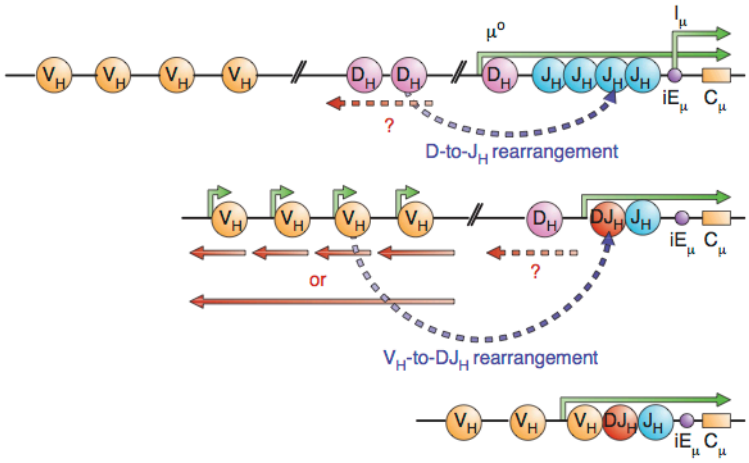


Figure 6. Schematic representation of the mouse *IgH* locus transcription and rearrangement. Top panel depicts events in early pro-B cell stage. Middle panel depicts events in late pro-B cell stage. Bottom panel depicts events in pre-B and mature B cell stage. Sense transcripts are depicted as green arrows. Rearrangement of gene segments is depicted as broken blue arrows. Antisense VH germline transcripts are depicted as red arrows. Antisense DH germline transcripts are depicted as broken red arrows. From Giallourakis et al., 2004.

Another correlate of loci accessibility is epigenetic regulation of the chromatin. Histone acetylation, a hallmark of transcriptionally and recombinationally active chromatin, appears to be activated in a stepwise process that coincides with the order of V(D)J recombination. In pro-B cells before D_H-to-J_H recombination occurs, the genomic region spanning the D_H, J_H, and Cμ exons becomes accessible to RAG as evident by H3 and H4 acetylation (Subrahmanyam and Sen, 2010). On the contrary, the V_H region remains hypoacetylated and is inaccessible to the recombination machinery (Subrahmanyam and Sen, 2010).

Trimethylation of the H3K4 residue (H3K4me3) has been shown to be highly enriched around active promoters (Heintzman et al., 2007). The functional consequence of this enrichment links histone modifications to the function of the RAG endonuclease. The RAG2 protein consists of core and noncore regions. The core RAG2 fragment facilitates cleavage of RSSs during recombination whereas the noncore fragment is hypothesized to facilitate faithful

recombinations of gene segments (Jones and Simkus, 2009). The RAG2 noncore fragment contains a non-canonical plant homeodomain (PHD) that has been shown to mediate interactions with chromatins (Matthews et al., 2007). In the context of V(D)J recombination, the recognition and binding to H3K4me3 by the RAG2 plant homeodomain (PHD) are postulated to promote V(D)J recombination by recruiting or retaining the RAG complex at antigen receptor loci (Matthews and Oettinger, 2009).

In addition to iEμ, a number of regulatory elements have been identified in the murine *IgH* locus. Recently, a 4kb intergenic region between $D_{FL16.1}$ and V_{H81X}, named intergenic control region 1 (IGCR1), has been shown to play a critical role in regulating *IgH* locus recombination (Guo et al., 2011). The IGCR1 contains two binding sites for CCCTC-binding factor (CTCF). CTCF is a highly conserved, ubiquitously expressed transcription factor that binds a GC-rich consensus sequence (Phillips and Corces, 2009). CTCF has been shown to exhibit transcriptional activation as well as transcriptional insulation activity via mediation of chromatin looping (Shih and Krangel, 2013). Mutation of these CTCF-binding elements (CBEs) increase transcription level of proximal V_H genes, resulting in decreased utilization of the distal V_H genes and loss of ordered V_H-to-DJ_H rearrangement (Guo et al., 2011). Moreover, lineage specificity was also affected, as V_H-to-DJ_H recombination was detected in CBE-mutant thymocytes (Guo et al., 2011). Therefore, IGCR1 has been proposed to insulate proximal V_H segments from rearranged D-J_H substrates, favoring recombination to more distal VHs (Figure 7).

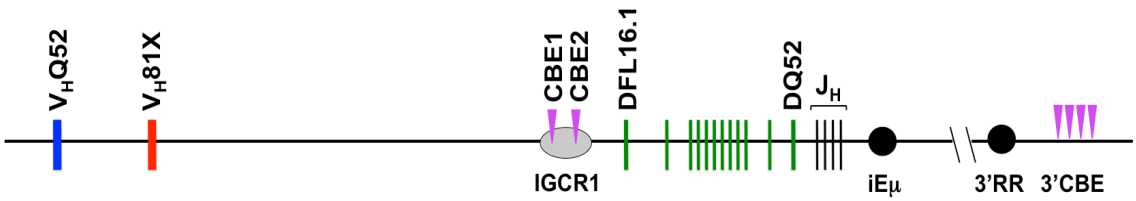


Figure 7. Schematic representation of the mouse IGCR1 organization. The most distal V_H segment is depicted as a blue line. The most proximal V_H segment is depicted as a red line. CBE sites are depicted as purple triangles. The D_H segments are depicted as green lines with the most 5' and 3' segment indicated. The J_H segments are depicted as black lines. Regulatory elements are depicted as black circles. From Guo et al., 2011.

Another major challenge during V(D)J recombination is how to bring together gene segments that reside over very large linear distances. A complex and dynamic modulation of antigen receptor loci conformation has been recently proposed to be responsible for such process. Several studies suggest a model where the *IgH* locus in pre-pro B cells is organized in multiple large "rosette-like" compartments or subdomains, each composed of multiple DNA loops (Subrahmanyam and Sen, 2012; Bossen et al., 2012). Upon the initiation of V(D)J recombination, these subdomains undergo further interactions required to bring the distal V_Hs segments in close proximity to the D_H and J_H segments for synapsis and cleavage by RAG.

The newly identified IgH regulatory elements termed Pax5-activated intergenic repeats (PAIRs) are believed to modulate the *IgH* locus contraction (Ebert et al., 2011). PAIRs are interspersed within the distal V_H gene segments, contain binding sites for a B cell-specific transcription factor E2A, and also interact with CTCF (Ebert et al., 2011). PAIRs have also been shown to induce pro-B cell-specific germline transcription and active chromatin modifications coinciding with the *IgH* locus contraction (Ebert et al., 2011). Further genetic studies will be needed to better elucidate the exact role of PAIRs in V(D)J recombination. Such locus

contraction can be visualized by three-dimensional FISH: *IgH* and *Igκ* loci undergo contraction specifically in rearranging cells, while they are in a decontracted state in other developmental stages and cell types (Hewitt et al., 2010; Roldan et al., 2005). A similar stage-specific contraction has been described at TCRα/δ and TCRβ loci (Shih and Krangel, 2013).

1.3. Secondary Antibody Diversification

Successful rearrangement of the IgH and IgL by V(D)J recombination results in the expression of IgM and/or IgD molecules on the surface of B lymphocytes. The expression of IgM allows B lymphocytes to migrate out of the bone marrow into peripheral lymphoid organs such as the spleen, lymph nodes, and gut-associated lymphoid tissue. At these sites, B cells encounter antigen and are stimulated to undergo CSR and SHM.

Class switching refers to a recombination event in which the constant region of the Ig molecule is exchanged for another class of constant region without altering antigen-binding specificity of the molecule (Keim et al., 2013). CSR occurs in mature B lymphocytes and unlike V(D)J recombination, this process does not occur in T lymphocytes. During CSR, the Cμ constant region is replaced by downstream CH exons such as Cγ, Cε, and Cα, resulting in a change in the antibody class from IgM to IgG, IgE, or IgA, respectively (Chaudhuri et al., 2007) (Figure 8). The recombination takes place between S regions, which lie just upstream of the various CH genes (Boboila et al., 2012). Each S region is preceded by its own transcriptional promoter that is responsive to certain cytokines and activators from the germinal center environment (Zhang et al., 2010). Transcription through the S region directs CSR to specific downstream C_H genes. This regulation ensures that antibodies with the most suitable effector function are produced (Chaudhuri et al., 2007). To complete CSR, DSBs in the donor Sμ region and a downstream acceptor S region are joined and the intervening region between the two S-region DSBs is deleted (Boboila et al., 2012). The synapsis and joining are mediated by general DNA end-joining and DNA repair pathways (Zarrin et al., 2007). The juxtaposition of the downstream C_H

gene to the V(D)J exon results in the production of a new constant region under the control of the V(D)J exon promoter (Chaudhuri et al., 2007) (Figure 8).

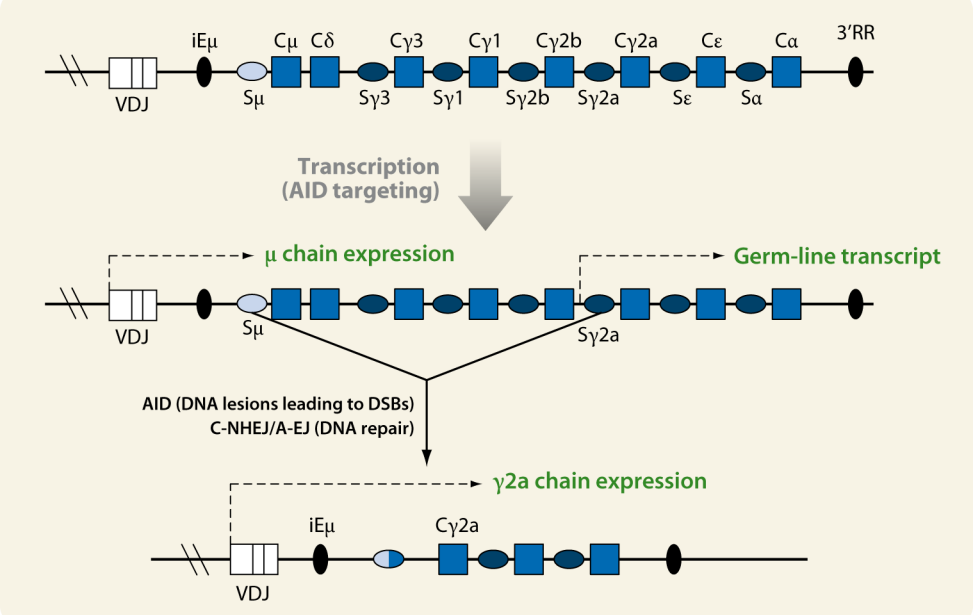


Figure 8. Schematic representation of the mouse *IgH* constant region locus and of the recombination events occurring in CSR. C_H genes are depicted as blue rectangles. S regions are depicted as blue ovals. Enhancer elements are depicted as black ovals. From Gostissa et al., 2011.

Mature B lymphocytes can also alter their antigen-binding specificity via SHM (Keim et al., 2013). SHM introduces a high rate of mutation via point mutations, small deletions, and insertions into the assembled variable region of Ig molecules, starting approximately 150 basepairs downstream of the IgV promoter and extending over 2 kb (Wang, 2013; Neuberger and Rada, 2006). The mutation frequency is estimated to be approximately 10^{-3} to 10^{-4} per basepair per cell division, can occur at either C/G or A/T pairs, and can be either transitions or transversions (Neuberger and Rada, 2006). These mutations allow for the selection of B lymphocytes with higher antigen-binding affinity BCR (Keim et al., 2013). The combination of CSR and SHM helps the immune system mount a more effective immune response.

CSR and SHM are initiated by activation-induced deaminase (AID) (Muramatsu et al., 2000). AID functions as a single-stranded (ss)-specific DNA deaminase. It catalyzes the dC-to-dU deamination in transcribed S regions and variable gene segments during CSR and SHM, respectively (Chaudhuri et al., 2007). AID preferentially recognizes RGYW and WRCY motifs (where R = purine, Y = pyrimidine, and W = A or T nucleotide) and canonical palindromic motif AGCT, which are enriched in S regions and within variable region exons (Chaudhuri et al., 2007). AID is recruited to ssDNA by transcription. Transcription through the S region results in the formation of RNA/DNA hybrids in which the non-template DNA strand can assume single-stranded conformation termed R-loop. These R loops provide ssDNA substrates for AID (Voung and Chaudhuri, 2012). The V regions, unlike the S region, do not contain the G:C-rich sequences needed for R-loop formation. Therefore, to complete SHM, AID interacts with ssDNA binding protein Replication Protein A (RPA). The AID:RPA complex has been demonstrated to bind to and deaminate transcribed DNA *in vitro* (Chaudhuri et al., 2004). AID does not generate mutation or DSBs directly, instead it generates a U:G mismatch DNA lesion. The processing of the dU:dG mismatch by the base-excision repair (BER) or mismatch repair (MMR) pathways result in DNA nicks that lead to mutations or DSBs (Di Noia and Neuberger, 2007). In the case of CSR, the DSBs in the donor and acceptor S regions are brought together for joining via ATM-dependent DSB response and C-NHEJ (Boboila et al., 2012). In the absence of C-NHEJ factors, CSR can still be carried out by alternative end-joining pathways at up to 50% of normal level (Yan et al., 2007). The A-EJ pathways will be discussed in greater detail in section 1.4.2. Experiments in *in vitro* stimulated B cells showed that AID activity in S regions is very robust and

can generate sufficient DSBs in the *IgH* locus such that over 50% of the cells can be induced to undergo CSR (Boboila et al., 2010).

Multiple transcriptional and post-transcriptional regulatory mechanisms control AID activity (Keim et al., 2013). Enhancers and promoter elements have been shown to restrict the transcription level of the *aicda* gene that encodes AID in a cell lineage- and developmental stage-specific manner (Voung and Chaudhuri, 2012). At least two microRNAs (miR-155 and miR-181b) post-transcriptionally downregulate AID mRNA and protein levels, leading to reduced CSR (de Yebenes et al., 2008; Dorsett et al. 2008; Teng et al. 2008). Moreover, post-translational modifications such as phosphorylation can also regulate AID activity. AID can be phosphorylated at multiple residues, resulting in an increased or reduced level of CSR and SHM (Voung and Chaudhuri, 2012). Lastly, although AID exhibits its activity on DNA in the nucleus, most AID in B cells is found in the cytoplasm, suggesting that AID activity may also be regulated by its localization (Geisberger et al., 2009).

1.4. The Role Of DNA Repair Pathways In V(D)J Recombination And CSR

V(D)J recombination and CSR generate DSBs in the genome that can lead to undesirable chromosomal aberrations. To maintain genomic integrity, mammalian cells have evolved two major DNA repair pathways: homologous recombination (HR) and classical non-homologous end joining (C-NHEJ) (Grabarz et al., 2012). HR is initiated by single-strand DNA (ssDNA) resection and repairs DSBs with high fidelity using a stretch of homologous sequence as a template, usually from the sister chromatid after DNA replication (Jasin and Rothstein, 2013). For this reason, HR mostly functions in post-replication repair during the S/G2 phase of the cell cycle (Grabraz et al., 2012). Unlike HR, C-NHEJ can join DNA ends without sequence homology such as in direct joining, or ends with short micro-homologies (MHs) (Bunting and Nussenzweig, 2013), and often requires end processing to prepare broken ends for ligation, therefore resulting in imprecise joins (Boboila et al., 2012). C-NHEJ is functional in all phases of the cell cycle and is the dominant pathway during the G0/G1 phase when HR is not active (Saintigny et al., 2007).

1.4.1. Classical Non-Homologous End Joining

The first step in the C-NHEJ repair reaction requires the binding of Ku70/Ku80 heterodimer to DSBs, where they may serve a protective function and also recruit other repair factors (Gu et al., 1997; Goedecke et al., 1999). Ku70/Ku80 are absolutely required for subsequent completion of the ligation reaction by XRCC4 and DNA ligase4 (Lig4). For this reason, Ku70/Ku80, XRCC4 and Lig4 are considered “core” C-NHEJ factors (Boboila et al., 2012) which are sufficient for ligation of DNA ends that do not require processing, such as signal ends

generated by RAG during V(D)J recombination. For the joining of other types of DNA ends, such as hairpin-sealed coding ends, further processing is required: binding of Ku70/Ku80 to DSBs activates the serine/threonine protein kinase, DNA-PKcs, which in turn phosphorylates and activates the Artemis nuclease (Helmink and Sleckman, 2012), necessary to cleave the hairpins to generate DNA ends suitable for ligation (Ma et al., 2002). Therefore, in the context of V(D)J recombination, C-NHEJ contributes to the diversity of the primary antibody and TCR repertoires via the imprecise joining resulting from such end-processing, combined with the addition of "N" (nontemplated) nucleotides by the template-independent DNA polymerase TdT (terminal deoxinucleotdyl transferase) (Lieber, 2010).

Another recently identified factor, Cernunnos/XLF, has been proposed to function in C-NHEJ, but its role is not yet fully clarified (Li et al., 2008; Zha et al., 2007; Ahnesorg et al., 2006; Buck et al., 2006). XLF directly interacts with the XRCC4-Ligase 4 complex (Ahnesorg et al., 2006; Buck et al., 2006; Callebaut et al., 2006) and it is believed to be involved in end-ligation and joining of DSBs with incompatible or blunt ends (Tsai et al., 2007). However, XLF-deficient mice to not show V(D)J recombination defects (Li et al., 2008), unless the ATM kinase (see below) is simultaneously deleted, suggesting ATM or other downstream factors may play a compensatory role for XLF deficiency in lymphocytes (Zha et al., 2011) (Figure 9).

C-NHEJ plays an important role in maintaining genome integrity by suppressing chromosomal translocation via directing DSBs to join within a chromosome as opposed to joining interchromosomally (Zhang et al., 2010). Moreover, it is absolutely required for V(D)J recombination, because RAG2 channels the repair of V(D)J breaks to this repair pathway, thereby excluding other pathways (Corneo et al., 2007). Accordingly, mice deficient for core C-

NHEJ factors show severe phenotypes, including immunodeficiency, growth retardation (in the case of Ku deficiency) or embryonic lethality due to neuronal apoptosis (in the case of XRCC4 or Lig4 deficiency), and increased sensitivity to DNA damage (Frank et al., 1998; Gu et al., 1997; Li et al., 1995; Taccioli et al., 1994).

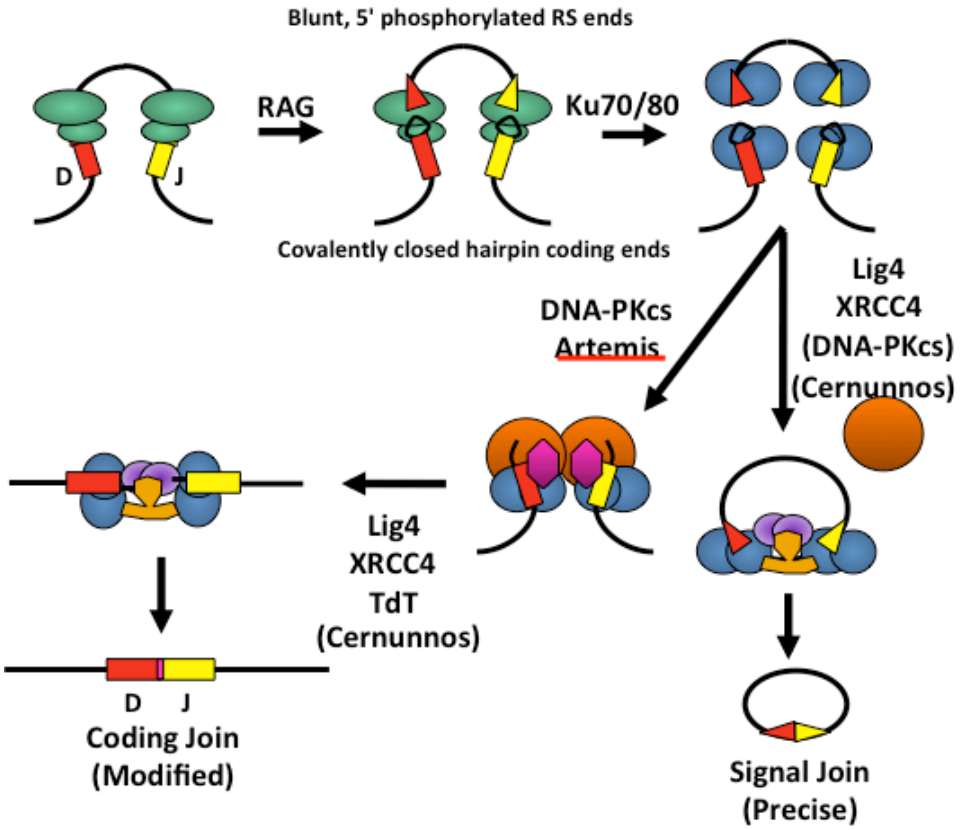


Figure 9. Schematic representation of joining of RAG-initiated DSBs by C-NHEJ. Adapted from Zhang et al., 2010

1.4.2. Alternative End Joining

Even though V(D)J recombination is dependent on C-NHEJ, joining of other types of DSBs, including those involved in chromosomal translocations (Zhu et al., 2002) and those generated during CSR (Soulas-Sprauel et al., 2007; Yan et al., 2007) can still take place in the absence of C-NHEJ, suggesting the existence of alternative end joining (A-EJ) pathways (Boboila

et al., 2012). In this regard, mature B cells deficient for core C-NHEJ factors still undergo CSR in culture at levels up to 50% of WT, indicating the robustness of A-EJ (Yan et al., 2007). Sequence analysis of CSR junctions from these mice revealed that A-EJ preferentially repairs broken DNA ends via MH-mediated mechanisms (Boboila et al., 2012). However, MH is not a requirement for A-EJ. In this regard, direct joining of I-SceI-initiated as well as AID-initiated DSBs has been detected in Ku80-deficient and Ku70- or dual Ku70/Lig4-deficient cells, respectively (Boboila et al., 2010; Guirouilh-Barbat et al., 2007, 2004). Many DNA repair factors have been implicated in A-EJ, including Parp1, Mre11, CtIP, XRCC1, Ligase1, and Ligase3; however, their exact roles remain to be elucidated (Boboila et al., 2012). Moreover, several evidences point to the existence of at least two separate A-EJ pathways (Boboila et al., 2012). Therefore, to date, A-EJ is best defined as DNA end joining that takes place in the absence of C-NHEJ factors. A-EJ is also regarded as a "translocation-prone" pathways, mostly based on the observation of increased MH usage at translocations junctions (Boboila et al., 2002). In this regards, C-NHEJ can prevent such aberrations by eliminating unrepaired DSBs or suppressing activity of A-EJ, or both (Gostissa et al., 2011).

1.4.3. The ATM-Dependent DNA DSB Response

Efficient repair of chromosomal DSBs requires not only the above-mentioned end joining activities, but also the function of a series of proteins collectively referred to as the DNA damage response (DDR) factors. One of the initial steps of the DDR is activation of the ATM kinase, a member of the PI3-kinase-related kinases (PI3KKs), which in turn phosphorylates many downstream substrates, including H2AX, 53BP1, MDC1, and RNF8 (Puebla-Osorio and

Zhu, 2008). Upon phosphorylation, these factors associate to form large macromolecular complexes, called foci, on the chromatin surrounding the site of damage, where they tether broken DNA ends and recruit repair factors (Alt et al., 2013; Bassing and Alt, 2004a). In addition, the activation of ATM initiates cell cycle checkpoints to arrest cell cycle progression while repair takes place (Alt et al., 2013) (Figure 10).

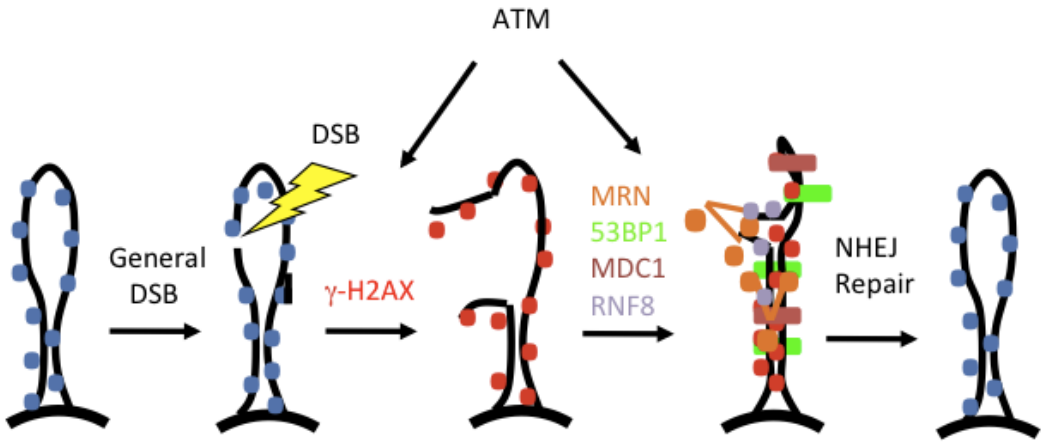


Figure 10. Schematic representation of ATM-dependent DNA damage response pathway. Adapted from Bassing et al., 2004.

ATM is present as an inactive homodimer that is recruited to sites of DSB by the Mre11/Rad50/Nbs1 (MRN) complex. Upon binding to MRN, ATM becomes activated via an autophosphorylation process that converts it into active monomers (Helmink and Sleckman, 2012). One of the first DDR responses initiated by activated ATM is the phosphorylation of the histone H2A variant H2AX; once phosphorylated on serine 139, H2AX is referred to as γ-H2AX and serves as a docking site to which other DDR factors, such as 53BP1 and MDC1 bind (Fernandez-Capetillo et al., 2004). Additionally, γ-H2AX is proposed to stabilize broken DNA ends in repair complexes, thereby facilitating effective DNA repair (Bassing and Alt, 2004b).

MDC1 is best characterized as an adaptor protein that recruits other repair factors to the site of DNA breaks so that they can be activated by ATM-dependent phosphorylation (Coster and Goldberg, 2010). This process results in an accumulation as well as an amplification of the DDR signal (Giunta and Jackson, 2011). MDC1 binds directly with γ-H2AX and ATM, resulting in the generation of a positive feedback loop to promote DNA repair and ATM retention (Derheimer and Kastan, 2011;). RNF8, a RING-finger E3 Ubiquitin ligase, is recruited to the repair complex upon phosphorylation of MDC1 (Yan and Jetten, 2008). RNF8 mediates ubiquitylation of H2AX, which is required for remodeling of chromatin and downstream recruitment and retention of 53BP1 at DSB sites (Mailand et al., 2007). The important role of DDR factors in promoting DNA repair is indicated by the fact that deficiency in many of them (including ATM, H2AX and MDC1) results in CSR defects and increased genomic instability in murine B cells (Franco et al., 2006; Ramiro et al., 2006). Strikingly, 53BP1 deficiency, while leading to the most severe reduction in CSR, only modestly affects general DNA breaks and translocations, suggesting still unidentified specific functions of this factor in class switching (Reina-San-Martin et al., 2007; Franco et al., 2006; Ward et al., 2004; Manis et al., 2004**)**.

In the context of V(D)J recombination, ATM plays a pivotal role in repairing RAG-initiated DSBs. ATM has been shown to promote the stability of coding ends in postcleavage complexes as coding ends are lost from the postcleavage complexes in ATM-deficient cells (Huang et al., 2007; Matei et al., 2007; Vacchio et al., 2007; Bredemeyer et al., 2006). Such stabilization of DSBs occurs through the phosphorylation of downstream substrates of ATM, all of which play a role in tethering DNA ends as discussed above. In the absence of ATM, the dissociation of these RAG-initiated DSBs allows them to join aberrantly to other unrepaired

DSBs, resulting in chromosomal deletions, inversions, and translocations (Mahowald et al., 2009; Bredemeyer et al., 2006).

1.4.4. The p53-dependent cell cycle checkpoint

ATM also initiates cell cycle checkpoints, mostly through activating tumor suppressor p53. Phosphorylation of p53 directly by ATM and indirectly by downstream kinases Chk1 and Chk2, activates the protein through the stabilization and increase of its transactivation activity (Roos and Kaina, 2013). Upon activation, p53 acts as a potent transcriptional activator or repressor that initiates a signaling cascade to prevent cell cycle progression and regulate pro- and anti-apoptotic genes, ultimately leading to the elimination of cells with persistent unrepaired DSBs (Vousden and Prives, 2009). The gene that encodes p53, *TP53*, is the most frequently mutated gene found in human cancers, highlighting the crucial role of this protein in safeguarding genome integrity (Meek, 2009). The role of p53 in tumor suppression is also demonstrated in mouse models. p53-deficient mice are prone to develop cancers, mostly thymic lymphomas while a small percentage of these mice develop B lineage lymphoma as well as other types of cancers such as sarcoma (Jacks et al., 1994; Donehower et al., 1992). The majority of p53-deficient thymic lymphomas is characterized by the predominance of aneuploidy and the lack of translocation in *Tcr* loci (Liao et al., 1998). A recent study by Gostissa and colleagues revealed that inactivation of p53 in mature B cells in mice led to the development of IgM-positive peripheral B cell lymphomas with clonal translocations that did not involve Ig loci or the c-myc oncogene (Gostissa et al., 2013). These translocations were not

recurrent and each tumor harbored different translocations, suggesting that the tumors did not arise from a common oncogenic event (Gostissa et al., 2013).

1.5. Chromosomal Aberrations and Cancer

Genomic rearrangements in cancer stem from alterations of chromosomes. These alterations include deletions, insertions, amplifications, duplications, inversions and translocations (Chen et al., 2010, Zhang et al., 2010). Chromosomal translocations result from the joining of large segments of DNA belonging to two non-homologous chromosomes (Zhang et al., 2010), and from faulty repair of DSBs generated through various mechanisms (**see below**). There are two types of chromosomal translocations: reciprocal and nonreciprocal. Reciprocal translocations take place when genetic material is exchanged between two chromosomes. They can be balanced, if the overall copy number of any section of the genome is not affected, or unbalanced, if they lead to deletion or duplication of genomic regions at the chromosomal breakpoints (Zhang et al., 2010). Translocations can promote oncogenic transformation if they affect copy number or expression of genes involved in control of cell proliferation and survival such as tumor suppressor genes and proto-oncogenes.

Recurrent translocations have been identified in lymphoid and solid tumors (Mitelman et al., 2007; Kuppers and Dalla-Favera, 2001). In the context of developing lymphocytes, the frequent DSBs generated at the *Ig* and *TCR* loci render these cells particularly susceptible to chromosomal aberrations. In humans, recurrent translocations fusing *Ig* or *TCR* loci to an oncogene are the hallmarks of B and T cell lymphomas and leukemias, respectively (Gostissa et al., 2011). Such a fusion event can lead to the deregulation of oncogene expression in two ways: by putting it under the control of potent *IgH* or *TCR* transcriptional enhancers or regulatory regions (Gostissa et al., 2011, **see below**), or by activating the oncogenes through the generation of oncogenic chimeric proteins. For example, the T(9;22) recurrent

translocations, referred to as the Philadelphia chromosome, are detected in the majority of chronic myelogenous leukemia (CML) patients and in a subset of adult patients with acute lymphoblastic leukemia (ALL) (Zhang and Rowley, 2011). This T(9;22) recurrent translocations result in a novel fusion of the *BCR* and *ABL1* genes, leading to constitutively activated ABL kinase. The activated ABL kinase activates a downstream signaling cascade leading to enhanced cell proliferation (Jabbour and Kantarjian, 2012).

1.5.1. Factors that Influence Oncogenic Chromosomal Translocation in Lymphocytes

Oncogenic translocation is a multi-step process that can be influenced by numerous factors. The generation of a translocation requires that at least two DSBs occur in a cell at the same time, and these breaks should be in relatively close spatial proximity in order to be joined together. Moreover, DNA repair pathways must fail to correctly repair these DBSs, allowing them to persist and be erroneously joined. Recurrent translocations appearing in tumors are subjected to cellular selection, in that cells harboring the translocation have proliferative advantages over the cells that did not have the translocation.

1.5.2. Role of DSBs in Chromosomal Translocations

The level of DSBs present in cells at a given time is one of the mechanistic factors that influence translocation frequencies (Alt et al., 2013). DSBs can result from cell-intrinsic factors, such as (1) oxidative stress from cellular metabolism byproducts, or (2) mistakes made during DNA replication or transcription (Alt et al., 2013; Tsai and Lieber, 2010; Hoeijmakers, 2009). Several studies have also shown that particular DNA sequences such as chromosome fragile

sites, repetitive elements, and various forms of non-B DNA structure can serve as sources of DSBs in the genome (Zhao et al., 2010; Hoeijmakers, 2009). Additionally, cell-extrinsic factors, such as type II topoisomerase inhibitors therapy, X-rays, and ionizing radiation can induce DSBs (Tsai and Lieber, 2010).

In lymphoid cells, DSBs can result from programmed DSBs that occur during lymphocyte development, or upon antigenic stimulation in germinal centers. DSBs from these events give rise to chromosomal translocations that can lead to lymphoid malignancies. The majority of lymphomas are of B cell origin and involve cells from all developmental stages (Kuppers, 2005). Many of these malignancies harbor chromosomal translocations that involve *Ig* genes and oncogenic partner genes. Several lines of evidence indicate that breakpoints in the *Ig* loci result from RAG- or AID-initiated DSBs during V(D)J recombination or CSR, respectively (Nussenzweig and Nussenzweig, 2010) while recurrent translocations observed in T cell leukemias or lymphomas involved RAG-initiated DSBs in the *Tcr* loci (Larmonie et al., 2013; Zha et al., 2010). Even though the generation of DSBs in *Ig* or *TCR* loci are well characterized, the source of DSBs in partner oncogene loci are not completely understood. In many cases, these DSBs have been proposed to arise from off-target RAG or AID activity (**see 1.5.2.1. and 1.5.2.2.)** (Figure 11).

1.5.2.1. Role of RAG Activity in Translocations

RAG-initiates DSBs at the RSSs flanking coding segments during antigen receptor gene assembly in developing B or T lymphocytes (Schatz et al., 2008). Such DSBs have been implicated in initiating oncogenic translocations involving *IGH* and *TCR* loci in many human lymphoid malignancies (Kuppers, 2005). For example, in Burkitt's lymphomas (BL), the hallmark

T(8;14) translocation juxtaposes the *c-myc* oncogene to the *IgH* locus, resulting in deregulation of the oncogene (Janz, 2006). Studies in C-NEHJ/p53-deficient mouse models provide direct evidence for the role of RAG-initiated DSBs in oncogenic translocations and lymphoma development. These mice routinely develop pro-B cells lymphomas with translocation breakpoints that link the *IgH* JH region to sequences near *c-myc* (Gostissa et al., 2011). These *IgH/c-myc* translocations are also RAG-dependent as RAG-deficiency in these backgrounds abrogates pro-B cell development (Vanasse et al., 1999).

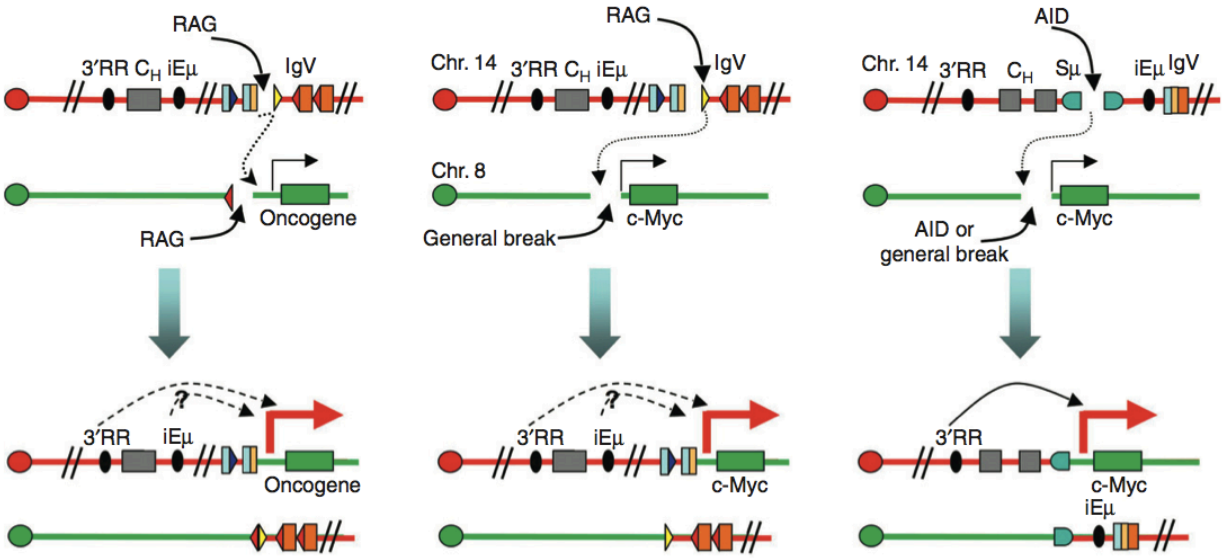


Figure 11. Schematic representation of chromosomal translocations involving aberrant V(D)J recombination and CSR. Left panel depicts aberrant V(D)J recombination between canonical RSS in the *IgH* locus and cryptic RSS near an oncogene. Middle panel depicts chromosomal translocation between RAG-initiated DSBs in the *IgH* locus to DSB generated by other mechanisms. The expression of the translocated oncogene is under the control of 3'RR or iEμ (depicted as broken arrows). Right panel depicts translocation involving aberrant joining of AID-initiated DSB in S region to DSBs generated by AID or other mechanisms, resulting in activation of the translocated oncogene by 3'RR (depicted as solid arrow). From Zhang et al., 2011.

Regarding the sources of DSBs in non-*Ig* or non-*Tcr* translocation partner loci, it is proposed that DSBs may result from off-target RAG activity that allows cleavage of cryptic RSSs outside of *Ig* or *Tcr* loci (Ji et al., 2010). Cryptic RSS is a 7-bp DNA sequence that shares the first

three nucleotides with a canonical heptamer sequence of RSS in *Ig* or *Tcr* loci (Zhang and Swanson, 2008). It is estimated that over 10 million cryptic RSSs might be present throughout the human genome (Lewis et al., 1997). A recent study also showed that changes in the 4th and 5th nucleotide from the consensus heptamer RS sequence result in premature release of signal ends from the postcleavage complex, leading to DSB repair by translocation-prone A-EJ and HR (Arnal et al., 2010). In addition to sequence-specific cleavage, RAG can also exhibit DNA-structure-dependent cleavage. Raghavan and colleagues showed that RAG can misrecognize and cleave non-B DNA structure in the major breakpoint region (Mbr) of the *BCL2* gene, suggesting that the translocations between *IgH* and *BCL2* in human follicular lymphoma (FL) arise as a result of misdirected V(D)J recombination (Raghavan et al., 2004).

A recent large-scale analysis of translocation breakpoints in T-ALLs demonstrates the role of RAG in translocations involving *TCR* loci. These aberrant translocations result in juxtaposition of oncogenes to *TCR* loci, thereby putting them under the control of *TCR* regulatory elements and enhancers (Dik et al., 2007). Two categories of *TCR* translocation mechanisms have been proposed. In “Type 1”, the translocations result from the joining of two RAG-initiated DSBs together. For example, in T-ALL patients harboring a *TCR-LMO2* translocation, the translocation junctions revealed a joining of *bona fide* RSS in *TCR* to cryptic RSS in *LMO2* (Larmonie et al., 2013). In “Type 2”, the most frequently detected type in human T-ALLs, the translocations result from the joining of RAG-initiated DSBs in antigen receptor loci or at cryptic RSSs in non-*TCR* loci to non-RAG-initiated DSBs (Larmonie et al., 2013). Both “Type 1” and “Type 2” translocations can also lead to interstitial deletions of genomic region such as the frequently observed *SIL-TAL* deletion (Larmonie et al., 2013). Strikingly, this study also

revealed that the majority of translocations involving the *TCR* loci occur in the *TCRD* locus, analogous to the involvement of murine *Tcrd* loci in ATM-deficient thymic lymphomas (Larmonie et al., 2013; Zha et al., 2010). Together, these results suggest that similar mechanisms may be responsible for oncogenic transformation in human T-ALL and ATM-deficient thymic lymphomas.

1.5.2.2. AID-Initiated Translocations

The DSBs generated as an intermediate for CSR, if unrepaired, can serve as a substrate for chromosomal translocation, as demonstrated by the accumulation of AID-dependent IgH locus breaks and translocations found in cultured B cells from mice deficient in C-NHEJ or DDR factors (Franco et al., 2006; Ramiro et al., 2006). AID has been shown to be required for *IgH/c-myc* translocations in primary B cells and mouse plasmacytomas (Casellas et al., 2008**).** Analysis of human lymphoid malignancies has revealed that recurrent oncogenic translocations often involve *IgH* S regions, such as the *IgH/c-myc* translocation found in sporadic Burkitt's lymphoma (BL) and the *IgH/bcl-6* translocation in diffuse large B cell lymphoma (DLBCL), among others (Kuppers, 2005).

The role of AID in generating DSBs at non-*Ig* translocation partners has been firmly established. Approximately 25% of the genes expressed in germinal center B cells are mutated by AID (Pasqualucci et al., 2001), some of which, such as *c-myc, Pim1, Pax5, and Rhoh* have been implicated in mouse and human lymphoid malignancies (Liu et al., 2008). Even though the mutation rates of these off-target sites are 100-fold lower than the *Ig* variable and S regions, cellular selection for oncogenic translocations during tumor development is believed to account

for recurrent translocation of the oncogenes (Gostissa et al., 2011; Liu et al., 2008). Moreover, AID activity has been shown to be necessary for *c-myc* locus breaks that participate in translocations with *IgH* in activated murine B cells (Robbiani et al., 2008) and overexpression of AID in p53-deficient mice lead to widespread genomic instability and development of B cell lymphomas with chromosomal translocations (Robbiani et al., 2009). Finally, recent studies of translocation patterns in murine activated B cells clearly showed the involvement of AID in the generation of DSBs genome-wide (Chiarle et al, 2011; Klein et al., 2011).

1.5.2.3. Partnership of RAG and AID in Translocations

RAG and AID are usually expressed in different B cell developmental stages. High levels of RAG are detected in developing B cells in the bone marrow whereas AID is mostly expressed in mature B cells responding to antigenic stimulation in the periphery. With functional DDR, RAG- and AID-initiated DSBs should not be present in the same cells since RAG-initiated DSBs would have been repaired before AID can induce DSBs. However, AID is also expressed at low levels in bone marrow B cells (Crouch et al., 2007; Feldhahn et al., 2007; Mao et al., 2004), and a recent investigation suggested that AID may facilitate off-target RAG activity in this context. According to this model, AID deamination activity at CpG islands would lead to C-to-T mutations, giving rise to T:G mismatches that serve as substrates for RAG-initiated DNA cleavage (Tsai et al., 2008). This model is supported by the breakpoint sequence analysis of human lymphoid malignancies, demonstrating that DSBs at the *bcl-2* major breakpoint region in FL and *bcl-1* major translocation cluster in mantle cell lymphoma were enriched around CpG sites (Tsai et al., 2008).

In the absence of ATM, unrepaired RAG-initiated DSBs have been shown to persist for multiple rounds of cell division and are detected in mature B cells in the periphery (Callen et al., 2007). In this context, cells with unrepaired V(D)J DSBs can escape cell cycle checkpoints, complete developmental stages, and harbor translocations that join RAG-initiated DSBs with AID-initiated DSBs in mature B cells. For example, AID-initiated DSBs at the *IgH* S regions and RAG-initiated DSBs at the *Igλ* locus have been reported to participate in translocations in mature B cells lacking the C-NHEJ factor XRCC4 (Wang et al., 2009). Furthermore, *IgH*/*Igλ* translocations are frequently found in mature B cell lymphomas arising in mice with mature B cell-specific inactivation of XRCC4 and p53 (Wang et al., 2008). These findings indicate that RAG is active in a subset of murine mature B cells, potentially as it drives the receptor editing process, and that some translocations found in human mature B cell tumors may similarly be ascribed to RAG and AID activity in the periphery.

1.5.3. Role of Spatial Genome Organization in Translocation Frequency

Recent studies of spatial genome organization reveal its role in the arrangement of chromosomal territories and impact on the formation of chromosomal translocations. In the interphase nucleus, chromosomes occupy discrete three-dimensional regions, known as chromosome territories (Cavalli and Misteli, 2013). These territories are arranged in non-random patterns: gene-rich chromosomes preferentially locate towards the nuclear interior and gene-poor chromosomes preferentially locate towards the nuclear periphery (Meaburn et al., 2007). There is a strong correlation between spatial proximity of chromosomes in interphase nuclei and translocation frequencies, whereby more proximal chromosomes

undergo translocation events more frequently than distantly located ones (Meaburn et al., 2007). A cytogenetic study by Parada and colleagues showed that translocations in ATM-deficient mouse lymphoma cell lines often involve chromosomes 12, 14, and 15, and these chromosomes are frequently found to be proximal to one another in normal mouse splenocytes (Parada et al., 2002). Other translocations such as T(9;22) involving BCR-ABL and and T(15;17) involving PML-RARα in human lymphocytes also exhibit the same pattern (Neves et al., 1999).

Similar cytogenetic studies with individual gene resolution showed a correlation between translocation frequency and translocation partner loci proximity. For example, in Burkitt's lymphoma, where *c-myc* undergoes translocation with *IgH*, *Igκ,* and *Igλ* at varying frequency, the distance between *c-myc* and its translocation partners corresponds with its frequency of translocation (Roix et al., 2003). Specifically, *c-myc* was more frequently colocalized with its preferred translocation partner *IgH*, as compared to *Igκ* and *Igλ* loci (Roix et al., 2003). Likewise, *IgH* frequently colocalized with *CCND1*, *bcl-2*, or *bcl-6* loci in mantle cell lymphoma (MCL), FL, and DLBCL, respectively (Roix et al., 2003).

Though proximity of loci and frequency of translocations are broadly correlated, both of these factors can also be tissue type- and cell type-specific. Studies of translocation frequency and spatial genome organization in different cell types revealed that tissue-specific translocation frequencies correlate with tissue-specific loci proximity (Parada et al., 2004). For example, chromosomes 12 and 15, which frequently translocate in B cell lymphomas, are proximal in lymphocytes but not in hepatocytes, whereas chromosomes 5 and 6, which often translocate in hepatomas, are proximal in hepatocytes, but not in lymphocytes (Parada et al., 2004).

Direct evidence for the influence of loci proximity in translocation frequency comes from Hi-C studies where chromatin interactions were analyzed in the absence of cellular selection to provide high-resolution maps of the genome-wide chromatin interactions (Lieberman-Aiden et al., 2009). These studies show that in the interphase nucleus of both cycling and G1-arrested mammalian cells, organization of chromosomal territories is such that intra-chromosomal interactions are largely favored over inter-chromosomal interactions (Zhang et al., 2012; Lieberman-Aiden et al., 2009), in particular at distances of about one megabase (Dixon et al., 2012). Accordingly, in G1-arrested ATM-deficient pro-B cell lines treated with IR to saturate the genome with DSBs, intra-chromosomal spatial genome organization causes regions along the entire length of an individual chromosome to become preferential partners for translocation with a fixed DSB, rather than regions on all other chromosomes (Zhang et al., 2012). These studies demonstrate that when DSB frequency at translocation partner loci is not a dominant factor (such as in the case when a large number of DSBs are induced by IR treatment in ATM-deficient pro-B cell lines or when DSBs are rare) spatial proximity of DSBs becomes an influential factor in determining genome-wide translocations (Alt et al., 2013; Zhang et al., 2012). These findings may have implications for the identification of mechanisms that control joining of random DSBs in tumor cells following chemo- or radiotherapy where numerous DSBs are the side effects of the treatment.

1.5.4. Roles of Antigen Receptor Locus Enhancers in the Selection of Oncogenic Translocation

Ig and *TCR* genes are regulated by an array of *cis*-acting promoters and enhancers. These elements help to regulate the ordered assembly of antigen receptor loci during

lymphocyte development and control the high-level expression of functionally assembled Ig and TCR loci in mature cells. These elements are also regarded as responsible for overexpression of oncogenes that lead to lymphomas. In the *Tcr* loci, major enhancer elements reside in various locations in each locus. In the *Tcrb* locus, the enhancer Eβ lies upstream of Cβ2. In the *Tcra/d* locus the Eδ lies between the Jδ and Jδ while the Eα lies downstream of the Cα (Dudley et al., 2005). Some of these enhancers have been shown to play a role in T cell tumor translocations and lymphoma development. For example, analysis of translocation junction in human T-ALLs revealed that approximately 28% of translocations involving *TCR* loci occur in *TCRB*, which corresponded with juxtaposition of Eβ to *TLX1*, resulting in overexpression of the oncogene (Larmonie et al., 2013; Le Noir et al., 2012)

Major known control elements in the *IgH* locus include the intronic enhancer (iEμ) and the *IgH* 3' regulatory region (*IgH* 3'RR), both of which have been linked to overexpression of oncogenes following translocations in mature B cell tumors such as BL, FL, and DLBCL. In transgenic mouse models of *c-myc* overexpression, iEμ leads to development of pre-B cell tumors in the majority of cases (Janz, 2006), while the *IgH* 3'RR promotes development of mature, Burkitt-like B cell lymphomas (Truffinet et al., 2007; Wang and Boxer, 2005). Moreover, deletion of the *IgH* 3'RR abrogates tumor development in a murine peripheral B cell tumor model based on inactivation of XRCC4 and p53 in mature B cells (Gostissa et al., 2009).

In humans, approximately 80% of BL cases can be traced to the T(8;14) translocation that juxtaposes *c-myc* to *IgH*, leading to deregulation of the oncogene (Magrath, 2012). Detailed analysis of the BL translocation breakpoints revealed two categories of junctions. In endemic BL, the *IgH* translocation breakpoints are located in the variable region upstream of

iEμ, resulting in the retention of both iEμ and I*gH*3'RR (Truffinet et al., 2007). In sporadic BL, breakpoints are located in the S regions, linking *c-myc* to the downstream portion of the *IgH* locus and retaining only the *IgH* 3'RR in the translocated allele (Hecht and Aster, 2000; Truffinet et al., 2007). These findings suggest the hypothesis that oncogenic translocations occur during SHM and CSR in endemic and sporadic BL, respectively. However, it is possible that endemic BL translocations take place during V(D)J recombination in progenitor or precursor B cells, but then remain inactive until later stages of development (Gostissa et al., 2011; Kuppers and Dalla-Favera, 2001). In this regard, *c-myc* breakpoints in endemic BL are usually located far upstream of the gene, at distances where iEμ may not be effective. These translocations may therefore become oncogenic only later in development, when the *IgH* 3'RR, which promotes transcription over large distances, becomes active (Gostissa et al., 2011). In line with this hypothesis, knock-in of *c-myc* between JH and iEμ did not lead to pro- or pre-B cell lymphomas, but only to mature B cell tumors (Janz, 2006), and pro-B cell tumors arising in C-NHEJ/p53-deficient mice routinely amplify c-myc (Difilippatonio et al., 2002; Zhu et al., 2002), suggesting that activity of iEμ alone is insufficient to activate *c-myc*.

Overall, these findings implicate the *IgH* 3' RR as a potent transcriptional element that can activate a high level of oncogene expression, and whose action may extend over hundreds of kilobases. Conversely, iEμ may have a more "local" and less potent activity that can still lead to transformation via deregulated expression of translocated genes, such as in the case of DLBCLs carrying *IgH*/*bcl-6* translocations, which retain iEμ but not the *IgH* 3'RR (Pasqualucci et al., 2003).

1.6. Role of DNA Repair in Suppressing Chromosomal Translocations

1.6.1. C-NHEJ/p53-deficient Mouse Models

Components of C-NHEJ pathway are essential for the repair of RAG-initiated DSBs during V(D)J recombination and have been implicated in maintaining genomic integrity via suppression of chromosomal rearrangements (Nussenzweig and Nussenzweig, 2010). Ku70, Ku80, XRCC4, and Lig4 are considered "core" C-NHEJ factors, which are absolutely required for all C-NHEJ reactions (Boboila et al., 2012). Ku70- or Ku80-deficient mice exhibit growth defects, increased sensitivity to DNA damaging agents, chromosomal instability, and severe combined immunodeficiency (SCID) due to an inability to carry out end-joining to complete V(D)J recombination (Gostissa et al., 2011). XRCC4 or Lig4 deficiency results in late embryonic lethality due to neuronal lethality; but developing embryos exhibit similar defects to those associated with Ku-deficiency including a block in B and T cell development due to a V(D)J recombination defect (Gao et al., 2000; Frank et al., 1998). On the contrary, mice deficient in DNA-PKcs or Artemis do not display growth defects and have "leaky" SCID phenotypes because they can still undergo a low level of V(D)J recombination (Gao et al., 1998; Rooney et al., 2003; Rooney et al., 2002; Taccioli et al., 1998).

Despite significant genomic instability observed in the absence of C-NHEJ, C-NHEJ deficient mice do not develop lymphoid or other neoplasia, most likely because C-NHEJ-deficient cells containing unrepaired G1 DSBs are actively eliminated via cell cycle checkpoints (Puebla-Osorio and Zhu, 2008). In this regard, combined deficiency for a C-NHEJ and p53 in mice leads to recurrent development of pro-B cell lymphomas (Gostissa et al., 2011). The hallmark of pro-B cell lymphomas in these backgrounds is the complex translocations with gene

amplification, referred to as complicons, involving the *IgH* locus (chromosome 12) and *c-myc* gene (chromosome 15) (Zhu et al., 2002). Detailed characterization of these tumors revealed clonal rearrangement of the JH gene segments with translocation breakpoints in the JH region joining to breakpoints located 70-700 kb downstream of *c-myc* locus on chromosome 15 (Zhu et al., 2002). This aberrant joining between chromosome 12 and 15 mediates the formation of dicentric chromosomes that ultimately lead to oncogenic amplification of the *c-myc* genes via breakage-fusion-bridge (BFB) mechanism, resulting in overexpression of *c-myc* and cellular transformation (Gostissa et al., 2011; Zhu et al., 2002). Subsequent studies further demonstrated that these tumors are RAG-dependent, since C-NHEJ/p53/RAG2-deficient mice did not succumb to pro-B cell lymphomas, instead they developed thymic lymphoma, as commonly observed in p53-deficient background (Vanasse et al., 1999). It is notable that inactivation of XRCC4 in p53-deficient mature B cells leads to surface Ig-negative lymphoma development with recurrent T(12;15) translocations that join *IgH* S region to region upstream of *c-myc*, resulting in ectopic activation of the oncogene (Gostissa et al., 2009; Wang et al., 2008).

1.6.2. C-NHEJ Deficiency in Human

Defects in C-NHEJ proteins in humans are associated with primary immunodeficiency and occasional development of lymphoid malignancy (O' Driscoll and Jeggo, 2006). For example, null mutations of the *Artemis* gene give rise to a SCID phenotype and the patients are unable to produce immunoglobulins (Moshous et al., 2001). Artemis-deficient patients suffer from pneumonitis, persistent viral diarrhea, growth defects, and radiosensitivity (Gennery et al., 2006). Hypomorphic *Artemis* mutations give rise to Omenn syndrome, a disorder typically

associated with hypomorphic RAG mutations, where patients present with a progressive combined immunodeficiency (CID) characterized by recurrent sinopulmonary or gastrointestinal infection, partial B- and T-cell deficiency, hypogammaglobulinemia, and some patients develop Epstein-Barr virus (EBV)-associated B cell lymphomas (Evans et al., 2006; Moshous et al., 2003).

Patients with hypomorphic mutations of *LIG4* present with variable symptoms, believed to result from varying residual mutant LIG4 activity. Some patients exhibit developmental delay, microcephaly, CID, and in one case, a predisposition to multiple myeloma (O' Driscoll and Jeggo, 2006; Roddam et al., 2002). Others had normal features and were not associated with immunodeficiency but exhibited cellular radiosensitivity following therapy for T-ALL (Dvorak and Cowan, 2011). Another rare form of human primary immunodeficiency is caused by mutations in Cernunnos/XLF gene. Clinical phenotypes of Cernunnos/XLF deficiency are similar to those observed in patients with LIG4 syndrome: CID, microcephaly, growth retardation, radiosensitivity, and recurrent bacterial and opportunistic infection (Buck et al., 2006). Some of the patients exhibited hypogammaglobulinemia, suggesting a role for Cernunnos/XLF in CSR (Buck et al., 2006).

1.6.3. DNA Damage Response Deficiency in Mouse and Human

ATM functions at the apex of the DNA damage response (DDR) pathway by promoting DSBs repair by C-NHEJ and mediating cell cycle checkpoints via activation of its downstream substrates (Bednarski and Sleckman, 2012). ATM deficiency in mice recapitulates many aspects of ataxia telangiectasia (AT) including increased predisposition to developing T cell lymphomas,

which will be discussed in greater detail in Section 6. Deficiency of downstream ATM substrates lead to unrepaired chromosomal breaks and translocation. However, unlike ATM-deficiency, other DDR-deficient mice showed little predisposition to lymphoid or other malignancies, unless p53 was also simultaneously inactivated. Genomic instability in DDR-deficient cells activates cell cycle checkpoints, resulting in the elimination of cells with unrepaired DSBs and translocations (Gostissa et al., 2011). Lymphoid malignancies in ATM-deficient background are reflective of the role of ATM in both DSB repair and activation of cell cycle checkpoints through the p53-dependent pathway.

For instance, H2AX-deficient mice exhibit increased genomic instability, growth retardation (Bassing et al., 2002; Celeste et al., 2002), and are prone to tumor development when also p53 in inactivated (Bassing et al., 2003; Celeste et al., 2003). H2AX-deficient mice have reduced numbers of lymphocytes and impaired CSR. Cytogenetic studies of H2AX-deficient B cells activated for CSR revealed that AID-initiated DSBs in the S region of the *IgH* locus were unrepaired and they accumulate as chromosomal breaks and translocations (Franco et al., 2006). Mice deficient for other DDR factors, including MDC1, 53BP1, or NBS1, also exhibit similar phenotypes of increased genomic instability and frequent unrepaired *IgH* chromosomal breaks upon activation for CSR (Franco et al., 2006; Lou et al., 2006; Ramiro et al., 2006; Reina-San-Martin et al., 2005).

In DDR-deficient background, p53-deficiency leads to increased B and T cell lymphoma incidence. For example, H2AX/p53-deficient mice succumb to either thymic lymphomas that do not involve *Tcr* loci or pro-B cell lymphoma with RAG-dependent clonal translocation similar to those observed in C-NHEJ/p53-deficient background (Bassing et al., 2003). 53BP1/p53-deficient

mice succumb to thymic lymphomas with occasional involvement of *Tcra/d* loci and infrequently develop B cell lymphomas (Morales et al., 2006; Ward et al., 2004).

Some DDR components are essential in normal development of an organism such that lack of functional proteins lead to embryonic lethality, as observed in the case of Mre11, Rad50, and Nbs1 knockout mice (Bohgaki et al., 2010). Nbs1 or Mre11 hypomorphic mutant mice exhibit increased radiosensitivity, defective cell cycle checkpoints, chromosome instability and immunodeficiency (Bohgaki et al., 2010). Rad50 hypomorphic mutants show partial embryonic lethality and exhibit progressive hematopoietic stem cell failure without increased radiosensitivity or defective cell cycle checkpoints (Bender et al., 2002).

Mutations of DDR components cause hereditary disorders in human. For example, *NBS1* mutations are associated with the Nijmegen breakage syndrome, a disorder characterized by microcephaly, radiosensitivity, growth defects, immunodeficiency, and increased predisposition to cancer (Gennery, 2006). Hypomorphic mutations of the *MRE11* gene result in ataxia telangiectasia like disorder (ATLD), characterized by cerebellar atrophy and radiosensitivity. Unlike AT, ATLD patients have almost normal immune responses, do not exhibit increased cancer predisposition or telangiectasia (Stewart et al., 1999).

1.7. Cancer Development and ATM Deficiency

1.7.1. Clinical Features of Ataxia-Telangiectasia

Ataxia-telangiectasia (AT) is a rare autosomal recessive disorder caused by mutations in the ataxia-telangiectasia mutated (ATM) gene. In human, ATM deficiency leads to cerebellar ataxia, ocular telangiectasia, immunodeficiency, increased sensitivity to ionizing radiation, and a predisposition to cancers (Perlman, 2012). Approximately 30% of AT patients develop cancer over their lifetime (McKinnon, 2012). Of these, lymphoid malignancies were found most frequently, representing a 200-fold increased likelihood compared with non-AT individuals (Matei et al., 2006). AT children primarily develop T cell acute lymphocytic leukemia (T-ALL) with an immature or mixed immature/mature phenotype while AT adults more commonly develop T cell prolymphocytic leukemia (T-PLL) or T cell chronic lymphocytic leukemia (T-CLL) with a more mature phenotype, suggesting a thymic and post-thymic cancer origin, respectively (Matei et al., 2006). Although rarer than T cell malignancies, B cell leukemias (B-ALL) and lymphomas (in particular non-Hodgkin's lymphoma, B-NHL) are also more frequent in AT patients than in the general population (Boultwood, 2001; Taylor et al, 1996).

1.7.2. Cytogenetic Aberrations in AT Lymphocytes

The role of ATM in maintaining genomic integrity is demonstrated by the presence of chromosomal alterations in nonmalignant cells from AT patients. Studies have shown that 10% of all T cells from AT patients harbor translocations involving chromosome 7 and 14, including inv(7)(p13;q35) and T(7;7)(p13;q35) that join *TCRB* to *TCRG*, T(7;14)(p13;q11) that join *TCRG* to *TCRA/D*, and T(7;14)(q35;q11) that join *TCRB* to *TCRA/D* (Taylor et al., 1996). These

translocations are believed to be non-oncogenic and are detected at much lower frequencies also in non-AT individuals (Taylor et al., 1996). Studies in nonmalignant B cells from adult AT patients revealed T(2;14) translocations involving *IGK* and *IGH* (Taylor et al., 1996; Metcalfe et al., 1991; Butterworth and Taylor, 1986; Kirsch et al., 1985). This presumably non-oncogenic translocation is specific to B cells, as it is not detected in T cells from the same patients (Butterworth and Taylor, 1986). Albeit less frequent, this type of translocations may arise from similar mechanisms as T cells translocation involving two *TCR* loci (Taylor et al., 1996). Based on recent findings in ATM-deficient murine lymphocytes, impaired V(D)J recombination due to ATM deficiency most likely account for these translocations (Huang et al., 2007; Matei et al., 2007; Vacchio et al., 2007; Bredemeyer et al., 2006).

1.7.3. Recurrent Translocations In ATM-Deficient Human Lymphoid Tumors

Although T-ALL is the most common malignancy in young AT patients, cytogenetic characterization of these tumors has been reported only for a very limited number of cases. Three studies from several years ago described patients carrying T(9;16)(q12;p13), T(7;14)(q35;q32), and T(14;14)(q11;q32) (Minegeshe et al., 1991; Vitolo et al., 1984; Wake et al., 1982). While exact translocation breakpoints have not been cloned, cytogenetic studies revealed that the T(7;14) and T(14;14) translocations involved the *TCRA/D* locus on chromosome 14q32 and the *TCRB* locus on chromosome 7q35 (Minegeshi et al., 1991; Vitolo et al., 1984; Wake et al., 1982). More recently, another report described amplification of sequences on chromosome 14q22 in a fourth AT T-ALL case, which correlates with recurrently amplified region in T cell lymphomas from ATM-deficient mice (see below) (Zha et al., 2010).

Translocations are better characterized in T-PLLs arising in adult AT patients, where common inv(14)(q11;q32), T(14;14)(q11;q32), and T(X;14)(q28;q11) involve the *TCRA/D* locus at 14q11 and oncogenes belonging to the *TCL1* family, *TCL1* and *1b* at 14q32, downstream of *IGH* and *MTCP1* at Xq28 (Taylor et al., 1996). Sequencing analysis of some of these translocation breakpoints identified the 14q11 breakpoint in the Jα region, suggesting involvement of V(D)J recombination breaks (Russo et al., 1989; Davey et al., 1988). These translocations are also commonly detected in sporadic T-PLL, suggesting shared mechanisms of oncogenic transformation between AT and non-AT T-PLL patients (Taylor et al., 1996; Stankovic et al., 2002). Of note, translocations between *TCRA/D* or *TCRB* and *TCL1* oncogenes can also be detected in normal lymphocytes from AT patients, albeit less frequently than translocations involving two *TCR* loci. Studies following the evolution of T cells bearing these aberrations over time revealed that they expand into large clones and acquire further mutations, and eventually develop into fully malignant tumors (Taylor et al., 1996).

Reflective of less frequent cases of B lineage lymphoma or leukemia in AT patients, a limited number of studies have analyzed chromosomal aberrations in this subset of patients. Sandlund and colleagues characterized one AT patient with Burkitt's lymphoma (BL) who lacked the classical T(8;14) translocation joining *MYC* to *IGH* normally found in BL patients (Sandlung et al., 2006). Instead, this patient harbored T(3;8) translocation joining *BCL6* to *MYC* with duplications of both oncogenes on other chromosomes (Sandlung et al., 2006). Given the limited number of studies, this finding does not rule out the involvement of immunoglobulin loci in the formation of translocations in AT B lineage cancers but highlight the need for further investigations of chromosomal abnormalities in ATM-deficient B cells.

1.7.4. ATM Mutations in Sporadic Lymphomas

Intensive efforts have been put into characterizing sporadic *ATM* mutations in neoplasia. Somatic ATM inactivation, often biallelic, has been reported in sporadic T- and B-ALL (up to 25% of cases), T-PLL (up tp 40% of cases), B cell chronic lymphocytic leukemia (B-CLL, up to 40% of cases) and adult B cell lymphomas, such as mantle cell lymphomas (MCL) (Gumy-Pause et al., 2004). Such ATM inactivation commonly results from deletion of chromosomal regions encompassing the *ATM* gene on human chromosome 11 or from missense or truncating mutations (Stankovic et al., 2002). Chromosomal translocations commonly found in AT T-ALL and T-PLL are detected in the sporadic forms of these tumors, suggesting shared mechanisms of oncogenic transformation between AT and non-AT cases (Taylor et al., 1996; Stankovic et al., 2002).

B-CLL is a heterogenous disease that result from the transformation of pre- or post-germinal center B cells. The origins of B cells that give rise to B-CLL impact clinical outcome of the disease as evident by better prognosis of B-CLL with somatic hypermutation than the subset lacking the mutation (Klein and Dalla-Favera, 2010). Of note, ATM mutations, and occasionally deletion, are specifically found in the subset of B-CLL cases that arise from pre-germinal center B cells, whereas B-CLL without ATM mutation derives from post-germinal center B cells, and usually carry mutations of other genes involved in tumor suppression, such as p53 (Stankovic et al., 2002). B-CLL has not been linked to recurrent translocations but is associated with genomic instability that give rise to chromosomal deletions and amplifications. These aberrations consequently lead to deregulated oncogene expression (via deletion of *DLEU-2* on 13q14), impaired DSB response (via deletion of *ATM* on 11q22), or deletion of tumor

suppressors (via deletion of *TP53* on 17p13) (Klein and Dalla-Favera, 2010). It is notable that deregulation of apoptotic pathways, such as Bcl2 overexpression, is also frequently observed in B-CLL (Guipaud et al., 2003).

ATM mutations are also frequently found in MCL, an aggressive non-Hodgkin's B cell lymphoma that usually derive from pre-germinal center B cells, although also a somatically hypermutated subgroup has been identified (Menanteau and Martinez-Climent, 2013). MCLs with ATM mutations display higher level of chromosomal aberrations than the tumors with wild-type ATM (Gumy-Pause et al., 2004). The signature translocation of MCL is T(11;14)(q13;32) which puts the cyclinD1 gene (*CCND1)* under the influence of IgH transcriptional control elements, leading to cyclinD1 overexpression and cellular transformation (Pileri and Falini, 2009). Approximately 20% of MCL cases show complex chromosomal rearrangements, including amplification of the *CCND1/IGH* translocation (Menanteau and Martinez-Climent, 2013) and secondary rearrangement or amplification of *c-myc* (Setoodeh et al., 2013). In some cases, high level amplification of other putative oncogenes has also been reported.

1.7.5. ATM Deficiency and Lymphoma Development in Mice

Several mouse models for AT have been generated with phenotypes similar to those observed in AT patients. ATM-deficient mice display growth retardation, neurologic dysfunction, and infertility (Barlow et al., 1996). Additionally, these mice are immunodeficient with severe defects in T cell maturation, hypersensitivity to gamma-irradiation, and highly predisposed to T cell lymphomas (Elson et al., 1996; Xu et al., 1996). T cell lymphoma in ATM-

deficient mice is of thymic origin that leads to enlargement of the thymus, resulting in fatality from heart or lung compression (unpublished observations). Approximately 12% of splenic T cells isolated from ATM-deficient mice harbor T(12;14) translocations involving the *Tcra/d* locus similar to nonmalignant T cells from AT patients (Liyanage et al., 2000). In contrast to the corresponding translocations found in AT patients, the majority of mouse T(12;14) translocation lead to deletion of *Tcl1* oncogene due to deletion of the telomeric portion of chromosome 12 (Zha et al., 2010; Liyanage et al., 2000). T cell lymphomas from ATM-deficient mice show translocation that join chromosome 14 to other chromosomes such as 12, 15, or X with T(12;14) found most commonly (Zha et al., 2010; Liyanage et al., 2000). All ATM-deficient tumors had either partial chromosome 15 duplications or trisomy 15, resulting in copy number increase of the *c-myc* gene (Zha et al., 2010; Liyanage et al., 2000).

Chromosome 14 translocations in ATM-deficient T cell lymphomas were initially believed to involve *Tcra,* as the tumors comprise of $CD4^+CD8^+$ DP thymocytes, a developmental stage where thymocytes undergo *Tcra* rearrangement (Callen et al., 2009; Liyanage et al., 2000). Moreover, such translocations were thought to confer cellular selection advantage via deregulation of oncogenes by the *Tcra* enhancer (Eα). On the contrary, a more recent study demonstrated that chromosome 14 translocations in ATM-deficient T cell lymphomas instead originated from RAG-initiated DSBs in the *Tcrd* locus as evident by the translocation breakpoint residing in the *Tcrd* locus and containing unrearranged Cδ region (Zha et al., 2010). Additionally, Eα is not required for the development of ATM-deficient T cells lymphomas since ATM-deficient mice lacking Eα still develop $CD4^+CD8^+$ T cell lymphomas similar to those arising in ATM-deficient mice, with recurrent T(12;14) translocations (Zha et al., 2010). In addition, detailed

analysis of these tumors showed a recurrent amplification of a 500-kb genomic sequence located approximately 10 Mb upstream of the *Tcra/d* locus on the residual centromeric fragment of chromosome 14, not the derivative T(12;14) (Zha et al., 2010). Although no known oncogene has been identified, the findings suggest that the amplification of this genomic region may be associated with the oncogenic cellular transformation of T cells in ATM-deficient background (Zha et al., 2010).

Zha and colleagues proposed that the mechanisms that lead to recurrent translocations and gene amplification in ATM-deficient background are analogous to the formation of complicon involving *IgH* and *c-myc* in C-NHEJ/p53-deficient pro-B cell lymphomas (Zhu et al., 2002). In the case of ATM deficiency, because of the combined defect in DSB repair and cell cycle checkpoints, RAG-initiated breaks in the *Tcrd* locus are not properly joined. These DSBs can persist in developing thymocytes and be propagated as dicentrics, either by replication through the original DSB (Zha et al., 2010; Zhu et al., 2002) or by joining to DSBs on the other chromosomes, including the other chromosome 14 (Zha et al., 2010). Dicentrics chromosomes are unstable during cell cycle and can subsequently undergo continuous breakage and rejoining in a BFB mechanism, thereby allowing for amplification of the region upstream of *Tcra/d* (Zha et al., 2010). It was proposed that the high frequency of T(12;14) may reflect (1) a high level of RAG-initiated DSBs on chromosome 12 that provide the most frequent substrate for translocation with chromosome 14, (2) the relative proximity of *IgH* to *Tcra/d* in developing thymocytes, or (3) cellular selection of T(12;14) for deletions of a tumor suppressor located in the telomeric portion of chromosome 12 (Zha et al., 2010). In this regards, analysis of T cell lymphomas in ATM-deficient background also revealed a deletion of *Bcl11b*, which lies in the

telomeric end of chromosome 12, a genomic region frequently deleted in these tumors, leading to hemizygous loss of this tumor suppressor gene (Zha et al., 2010). The loss of *Bcl11b* may confer growth advantage and impose cellular selection for cells harboring T(12;14) translocations. Of note, *BCL11B* was shown to be inactivated via deletion or missense mutations in a subset of human T-ALL, further demonstrating the role of Bcl11b as a tumor suppressor (Gutierrez et al., 2011).

1.7.6. The Link Between ATM-Deficient Mouse Models And Cancer Development In AT Patients

The involvement of the *Tcrd* locus in murine T cell lymphomas in ATM-deficient background indicate the possibility of the *TCRD* locus being the frequent translocation partner in AT patients with T-ALL or T-PLL. Even though the occurrence of this event is not yet established due to the very limited information available on translocations breakpoints in AT tumors, a recent characterization of translocation junctions in a large cohort of adult T-ALL showed that the majority (67%) of *TCR* translocations involve the *TCRD* rather than the *TCRA* locus (Larmonie et al., 2013). In these tumors, the status of ATM was not verified; however to date, there is no evidence indicating differences between T-ALL in AT compared with non-AT individuals, and indeed it is likely that the majority of T-ALL has mutated ATM alleles (Stankovic et al., 2002). Therefore, it would be expected that T-ALL developing in AT patients harbored translocation involving the *TCRD* locus as well.

ATM mutations can give rise to the same lymphoid malignancies whether it occurs in the germline or in somatic cells. However, the predominance of the cancer subtypes is different between inherited mutation or acquired mutation in hematopoietic cells. In this context, T cell

lymphomas/leukemia with recurrent translocation involving the *TCR* loci are found more frequently in AT patients as well as ATM-deficient mouse models. On the contrary, somatic *ATM* inactivation leads primarily to B cell tumor development. The study by Zha and colleagues is the first to propose mechanisms of T cell lymphoma development in ATM-deficient background. This study may provide insights into mechanisms that account for the amplification of the region upstream of chromosome 14q22 in human T cell lymphomas. Few investigations have been done in B cell tumors in AT patients and no B lineage malignancies have been observed in ATM-deficient mice. Therefore, the study described here demonstrates the first cases of mature B cell lymphomas in ATM-deficient background. Detailed characterization of this tumor may shed light on mechanisms of B cell lymphoma development in AT patients or patients with somatic ATM mutation.

CHAPTER 2: MATURE B CELL LYMPHOMA DEVELOPMENT IN ATM-DEFICIENT BACKGROUND

2.1. Summary

ATM senses DNA double-stranded breaks (DSBs) and facilitates their repair. ATM-deficiency predisposes to B and T cell lymphomas in humans but only to T cell lymphomas in mice. We now describe compound mutant mice that, when ATM deficient, consistently develop IgM+ mature B cell lymphomas. These tumors harbor oncogenic *c-myc* amplifications generated by a breakage-fusion-bridge mechanism (BFB) from dicentric chromosomes formed through fusion of V(D)J recombination-initiated *IgH* DSBs to sequences downstream of *c-myc*. Genome-wide cloning of *c-myc* DSB translocations in ATM-deficient IgM+ mature B cells activated by various methods revealed a 30 megabase region downstream of *IgH* to be a translocation hotspot due to frequent DSBs newly generated by BFB. We propose that ATM-deficiency allows V(D)J recombination DSBs within *IgH* in developing B cells to generate dicentric translocations that via BFB cycles generate new DSBs and *c-myc* gene amplifications that can contribute to oncogenic transformation of mature B cells.

2.2. Introduction

The B cell antigen receptor (BCR) and its secreted antibody form are heterodimers comprised of immunoglobulin (Ig) heavy (IgH) and light (IgL) chains (Jackson et al., 2013). T cell antigen receptors are similarly comprised of either αβ or γδ heterodimers (Cobb et al., 2006). The N-terminal portion of Ig and TCR chains contains the antigen binding variable (V) region. Exons that encode Ig or TCR variable regions are assembled in developing bone marrow (BM) B cells and developing thymic T cells by V(D)J recombination of germline V, D, and J segments

(Jackson et al., 2013). V(D)J recombination is initiated in progenitor B and T cells by the "RAG" endonuclease (Nishana and Raghavan, 2012). RAG initiates V(D)J recombination by introducing DNA double-strand breaks (DSBs) at borders of a pair of V, D, or J segments and short conserved flanking recombination signal sequences (RSSs) (Schatz et al., 2008). RAG-initiated DSBs at two participating V, D, or J segments occur in the G1 cell cycle phase where they are joined exclusively by the classical non-homologous DNA end-joining (C-NHEJ) DSB repair pathway (Helmink and Sleckman, 2012).

The C-terminal constant region of IgH chains can be encoded by several different sets of exons (CHs) that determine Ig class (e.g IgM, IgG, IgA) and effector function (Cobb et al., 2006). IgH V(D)J exons are assembled at the germline JH region just upstream of the Cμ exons, the first set of CH exons expressed. Transcription from the V(D)J exon is terminated downstream Cμ to generate an mRNA that encodes a μ IgH chain (Perlot and Alt, 2008). Subsequent assembly of an IgL variable region exon and expression of an IgL protein generates IgM, which is deposited on the surface of the resulting IgM+ B cells (Vettermann and Schlissel, 2010). Surface IgM-expressing B cells migrate from the BM to peripheral lymphoid organs, such as the spleen, where upon antigen stimulation they undergo IgH class switch recombination (CSR). CSR changes antibody effector functions by replacing Cμ exons with one of several sets of CH exons that lie 100-200kb downstream in the *IgH* locus (Keim et al., 2013). CSR employs DSBs initiated by activation-induced cytidine deaminase (AID) in large repetitive switch (S) regions that precede the various sets of CH exons (Chaudhuri et al., 2007). DSBs in the donor Sμ region are joined to the DSBs in an acceptor S region by end-joining to complete CSR (Chaudhuri et al., 2004).

Chromosomal translocations result from joining two separate DSBs together and are inter-chromosomal if they join DSBs on two different chromosomes (Zhang et al., 2010). Depending on how DSBs are joined, inter-chromosomal translocations can join the telomeric end of one chromosome to the centromeric end of another or they can result in dicentric chromosomes and/or acentric chromosome fragments (Bunting and Nussenzweig, 2013). Translocations also can also occur intra-chromosomally if they join two separate DSBs on the some chromosome. Such intrachromosomal translocations might result in deletions or inversions depending which ends of the DSBs are joined (Zhang et al., 2010). The high frequency of DSBs generated during B and T cell development by V(D)J recombination and during B cell activation for CSR provide substrates for chromosomal translocations (Gostissa et al., 2011). Indeed, spontaneous B or T cell lymphomas in humans and certain mouse lymphoma models often harbor recurrent translocations that fuse Ig or TCR loci to oncogenes and/or join them to other genomic DSBs in a way that deletes tumor suppressors (Mitelman et al., 2007).

C-NHEJ maintains genomic integrity by re-joining of DSBs and, thereby, suppressing chromosome breaks and translocations (Gostissa et al., 2011). C-NHEJ is important in G1 when homologous recombination, the other major DSB repair pathway, is not available (Boboila et al., 2012). Despite marked genomic instability, C-NHEJ-deficient mice do not routinely develop lymphoid tumors or other cancers, at least in part, due to elimination cells containing unrepaired G1 DSBs by the G1/S cell cycle checkpoint (Puebla-Osorio and Zhu, 2008). In this regard, combined deficiency in mice for C-NHEJ and p53, a tumor suppressor that activates the G1 checkpoint, leads to recurrent pro-B cell lymphomas in mice (Gostissa et al., 2011). A hallmark of these pro-B lymphomas is complex translocation with *c-myc* oncogene gene

amplifications, referred to as complicons (Zhu et al., 2002). Such complicons derive from fusion of RAG-initiated *IgH* DSBs on chromosome 12 to DSBs downstream of *c-myc* on chromosome 15, generating dicentric chromosomes that lead to oncogenic *c-myc* amplification via a breakage-fusion-bridge (BFB) mechanism (Zhu et al., 2002). A key aspect of the mechanism of complicon formation is the persistence of unrepaired RAG-initiated *IgH* DSBs due to (1) the absence of C-NHEJ and (2) impaired G1 checkpoint due to the absence of p53. Consequently, the unrepaired DSBs are replicated and can form the dicentrics that promote *c-myc* amplification (Zhu et al., 2002).

Unrepaired DSBs in G1, including V(D)J recombination and CSR associated DSBs, activate Ataxia Telangiectasia-mutated-kinase (ATM), which in turn activates downstream factors that form complexes in chromatin surrounding DSBs, preventing premature DSB separation and promoting joining by C-NHEJ (Bednarski and Sleckman, 2012). Additionally, this ATM-dependent DSB response also activates p53 to enforce the G1 checkpoint (Helmink and Sleckman, 2012). Human ATM-deficiency leads to Ataxia Telangiectasia (AT), a condition that includes immunodeficiency and increased predisposition to T and B cell cancers (Perlman et al., 2012). ATM inactivation in mice recapitulates several aspects of AT, including predisposition to thymic lymphoma, but does not predispose to B cell lymphoma (Jacks et al., 1994; Donehower et al., 1992). Murine ATM-deficient thymic lymphomas harbor recurrent translocations involving RAG-initiated DSBs in the TCRδ locus on chromosome 14 that lead to chromosome 14 dicentrics and complicons, as well joins of TCRδ to sequences spread over a many megabase regions downstream of *IgH* to create T(12;14) translocations that delete the Bcl11b haploinsufficient T cell tumor suppressor (Gutierrez et al., 2011; Zha et al., 2010). Thus, the chromosomal

aberrations of ATM-deficient thymic lymphomas appear mechanistically similar to those found in C-NHEJ/p53-deficient pro-B cell lymphomas, potentially because ATM deficiency leads to both a V(D)J joining defect and a G1-checkpoint defect. Notably, ATM-deficiency also allows unrepaired RAG-initiated *IgH* DSBs in developing pro-B cells to persist, via unknown mechanisms, as large centromeric chromosome 12 fragments into mature B cells where they can participate in chromosomal translocations (Callen et al., 2007).

We hypothesized that the frequency of B cells with *c-myc* DSBs available for translocation or their survival may be rate-limiting factors for B cell lymphoma development in the ATM-deficient background. We now describe ATM-deficient mature B cell lymphoma mouse models generated on the basis of these hypotheses. Our studies of these tumors suggest that ATM suppresses mature B cell lymphomas in susceptible backgrounds by preventing the conversion of unrepaired RAG-initiated *IgH* breaks in progenitor B cells from progressing into dicentric chromosomes that, through BFB cycles, give rise to gene amplifications that promote transformation of mature B cells.

2.3. Results

2.3.1. RAG Target Sequence In *c-myc* Leads To Mature B Cell Lymphoma In ATM-Deficient Mice

To test the hypothesis that DSB frequency in oncogenes such as *c-myc* may be a rate-limiting factor for B cell lymphoma development in ATM-deficient background, we generated ATM-deficient mice with a RAG target sequence ("DJβ" cassette) inserted in intron 1 of *c-myc*. This cassette is efficiently cut by RAG during B cell development (Ranganath et al., 2008), therefore providing a source of unrepaired DSBs that could serve as translocation substrates in the context of ATM-deficiency. To obtain the experimental cohort, we intercrossed $ATM^{+/-}$ mice that were either heterozygous or homozygous for the $c\text{-}myc^{DJ\beta}$ allele (see Material and Methods for details). The $ATM^{-/-}$ $c\text{-}myc^{DJ\beta/DJ\beta}$ and $ATM^{-/-}$ $c\text{-}myc^{DJ\beta/wt}$ offspring from these breedings (collectively referred to as "DA") was monitored for tumor development. $ATM^{-/-}$ mice lacking $c\text{-}myc^{DJ\beta}$ allele, $ATM^{+/-}$ $c\text{-}myc^{DJ\beta/DJ\beta}$, and $ATM^{+/-}$ $c\text{-}myc^{DJ\beta/wt}$ mice were kept as controls. Out of a total of 17 DA mice, 5 developed peripheral B cell lymphoma and 12 developed T cell lymphoma, based on surface marker analysis (see below). DA mice that succumbed to B cell lymphoma had enlarged peripheral lymph nodes, mesenteric lymph nodes, spleen, and thymus. Metastasis to the liver or kidney was also observed (**Table 1**). T cell lymphomas from DA mice as well as $ATM^{-/-}$ littermate controls showed the same tumor characteristics as those previously reported in ATM-deficient background (Zha et al., 2010) (**Table 2**). None of the $ATM^{+/-}$ $c\text{-}myc^{DJ\beta}$ littermate controls succumbed to lymphoma or other neoplasia (**Table 2**).

Table 1. Summary of tumor mice analyzed from DA cohort. Cause of death is listed in the third column. The genotype of c-myc$^{DJ_\beta}$ allele is indicated. Results from Southern blot analysis with 3'Myc, Cμ, JH, and Jκ probed are shown. Red indicates mice that succumbed to B linage lymphomas.

Mouse ID	Age (days)	COD	c-myc allele	*c-myc*	Cμ	JH	Jκ
DA 416	79	thymic lymphoma	DJ/DJ	N/D	N/D	N/D	N/D
DA 351	87	thymic lymphoma	DJ/DJ	N/D	N/D	N/D	N/D
DA 607	90	IgM+ lymphoma	DJ/DJ	Amp.	Not amp.	Rearr. not Amp.	Rearr.
DA 64	95	IgM+ lymphoma	DJ/DJ	Amp.	Amp.	Rearr. and Amp.	Rearr.
DA 360	96	IgM+ lymphoma	DJ/DJ	Amp.	Amp.	Rearr. and Amp.	Rearr.
DA 359	102	thymic lymphoma	DJ/DJ	N/D	N/D	N/D	N/D
DA 107	103	IgM+ lymphoma	DJ/Myc	Not amp.	Amp.	Rearr. and Amp.	Rearr.
DA 306	104	thymic lymphoma	DJ/Myc	Not amp.	N/D	N/D	N/D
DA 366	114	thymic lymphoma	DJ/DJ	N/D	N/D	N/D	N/D
DA 268	116	thymic lymphoma	DJ/Myc	N/D	N/D	N/D	N/D
DA 458	125	thymic lymphoma	DJ/DJ	N/D	N/D	N/D	N/D
DA 403	135	IgM+ lymphoma	DJ/DJ	Amp.	Amp.	Rearr. and Amp.	Rearr.
DA 367	141	thymic lymphoma	DJ/DJ	N/D	N/D	N/D	N/D
DA 32	160	thymic lymphoma	DJ/Myc	Not amp.	N/D	N/D	N/D
DA 352	189	thymic lymphoma	DJ/DJ	N/D	N/D	N/D	N/D
DA 22	207	thymic lymphoma	DJ/DJ	Not amp.	N/D	N/D	N/D
DA 21	220	thymic lymphoma	DJ/DJ	Not amp.	N/D	N/D	N/D

Amp. = Amplified signal
Rearr. = Rearrangement detected

Table 2. Control mice from DA cohort. Age at the time of sacrifice is indicated in the second column. Cause of death (COD) is indicated in the third column. Genotype of each mouse is indicated in the last column.

Mouse ID	Age (days)	COD	Genotype
288*	91	T cell lymphoma	ATM-/-
242*	99	T cell lymphoma	ATM-/-
66*	102	T cell lymphoma	ATM-/-
353*	118	T cell lymphoma	ATM-/-
487*	120	T cell lymphoma	ATM-/-
279*	147	T cell lymphoma	ATM-/-
115*	148	T cell lymphoma	ATM-/-
483*	153	T cell lymphoma	ATM-/-
43	239	sacrificed	ATM+/- DJβ/Myc
34	244	sacrificed	ATM+/- DJβ/DJβ
176	261	sacrificed	ATM+/- DJβ/DJβ
177	261	sacrificed	ATM+/- DJβ/DJβ
179	261	sacrificed	ATM+/- DJβ/DJβ
2	293	sacrificed	ATM+/- DJβ/Myc
3	293	sacrificed	ATM+/- DJβ/Myc
4	293	sacrificed	ATM+/- DJβ/Myc
5	293	sacrificed	ATM+/- DJβ/Myc
13	302	sacrificed	ATM+/- DJβ/Myc
14	302	sacrificed	ATM+/- DJβ/Myc
16	302	sacrificed	ATM+/- DJβ/Myc
9	311	sacrificed	ATM+/- DJβ/Myc
341	325	sacrificed	ATM+/- DJβ/Myc
335	389	sacrificed	ATM+/- DJβ/DJβ
339	389	sacrificed	ATM+/- DJβ/DJβ
12	393	sacrificed	ATM+/- DJβ/Myc
15	393	sacrificed	ATM+/- DJβ/Myc
349	>506	alive	ATM+/- DJβ/DJβ
337	>574	alive	ATM+/- DJβ/DJβ
338	>574	alive	ATM+/- DJβ/DJβ

* indicates $ATM^{-/-}$ from DA cohort

2.3.2. Bcl-2 Transgene Promotes Mature B Cell Lymphoma In ATM-Deficient Mice

Although all DA mice succumbed to lymphomas, the penetrance of B cell lymphomas was approximately 25%. We hypothesized that this may be due to poor survival of ATM-deficient B cells. To test this hypothesis, we seek to confer survival advantage to ATM-deficient B cells by enforcing expression of the anti-apoptotic factor Bcl-2. Eμ-Bcl-2 transgenic mice overexpress Bcl-2 in the B cell compartment, and develop B cell hyperplasia (Strasser et al., 1994, McDonnell et al., 1989). Mice harboring Eμ-Bcl-2 transgene were crossed with $ATM^{+/-}$ mice and the resulting $ATM^{+/-}$Eμ-Bcl-2 offspring were then bred together to obtain experimental $ATM^{-/-}$ Eμ-Bcl-2 (referred to as "AB") mice. $ATM^{-/-}$ mice lacking the Eμ-Bcl-2 transgene and $ATM^{+/-}$Eμ-Bcl-2 mice were kept as controls. Of eight AB mice analyzed, four developed B lineage lymphomas (**Table 3)**. Three of these mice exhibited the same tumor presentation as B cell lymphomas in DA cohort while AB36 developed a tumor mass in the gastrointestinal cavity with enlarged spleen, but normal size thymus and peripheral lymph nodes. The other four AB mice developed T cell lymphomas (**Table 3**), which again were similar to those obtained in $ATM^{-/-}$ deficient animals in this and previous studies (**Table 4** and Zha et al., 2010). None of the $ATM^{+/-}$ Eμ-Bcl-2 littermate controls succumbed to lymphomas or other neoplasia (**Table 4**). Strikingly, when we generated ATM-deficient mice that were either heterozygous or homozygous for the c-$myc^{DJβ}$ allele and carried the Eμ-Bcl-2 transgene (referred to as "DAB"), we obtained 100% penetrance of B cell lymphoma (**Table 5**). 7 out of 7 mice developed B cell lymphomas with the same characteristics as those found at lower frequency in DA and AB cohorts. $ATM^{+/-}$ c-$myc^{DJβ}$ Eμ-Bcl-2 or $ATM^{+/-}$ c-$myc^{DJβ}$ control mice did not develop lymphomas or other neoplasia (**Table 6**).

Table 3. Summary of tumor mice analyzed from AB cohort. Cause of death is listed in the third column. The genotype of c-myc$^{DJ_\beta}$ allele is indicated. Results from Southern blot analysis with 3'Myc, Cμ, JH, and Jκ probed are shown. Red indicates mice that succumbed to B linage lymphomas.

Mouse ID	Age (days)	COD	c-myc allele	*c-myc*	Cμ	JH	Jκ
AB 225	84	IgM- lymphoma	-	Amp.	Amp.	Rearr. and Amp.	deleted
AB 92	100	thymic lymphoma	-	Not amp.	N/D	N/D	N/D
AB 67	104	IgM+ lymphoma	-	Amp.	Amp.	Rearr. and Amp.	Rearr.
AB 117	113	thymic lymphoma	-	N/D	N/D	N/D	N/D
AB 143	116	IgM+ lymphoma	-	Amp.	Amp.	Rearr. and Amp.	Rearr.
AB 167	157	thymic lymphoma	-	N/D	N/D	N/D	N/D
AB 106	133	thymic lymphoma	-	N/D	N/D	N/D	N/D
AB 36	297	IgM- lymphoma	-	Not amp.	Not amp.	deleted	Rearr.

Amp. = Amplified signal
Rearr. = Rearrangement detected

Table 4. Control mice from AB cohort. Age at the time of sacrifice is indicated in the second column. Cause of death (COD) is indicated in the third column. Genotype of each mouse is indicated in the last column.

Mouse ID	Age (days)	COD	Genotype
116^	72	T cell lymphoma	ATM-/-
78^	106	T cell lymphoma	ATM-/-
60^	117	T cell lymphoma	ATM-/-
86^	153	T cell lymphoma	ATM-/-
55^	154	T cell lymphoma	ATM-/-
58^	155	T cell lymphoma	ATM-/-
101	247	sacrificed	ATM+/- Eμ-Bcl-2
90	337	sacrificed	ATM+/- Eμ-Bcl-2
42	447	sacrificed	ATM+/- Eμ-Bcl-2
33	467	sacrificed	ATM+/- Eμ-Bcl-2
34	467	sacrificed	ATM+/- Eμ-Bcl-2
20	471	sacrificed	ATM+/- Eμ-Bcl-2
31	686	sacrificed	ATM+/- Eμ-Bcl-2
24	757	sacrificed	ATM+/- Eμ-Bcl-2
21	804	sacrificed	ATM+/- Eμ-Bcl-2

^ indicated ATM$^{-/-}$ from AB cohort

Table 5. Summary of tumor mice analyzed from DAB cohort. Cause of death is listed in the third column. The genotype of c-myc$^{DJ_\beta}$ allele is indicated. Results from Southern blot analysis with 3'Myc, Cμ, JH, and Jκ probed are shown. Red indicates mice that succumbed to B linage lymphomas.

Mouse ID	Age (days)	COD	c-myc allele	*c-myc*	Cμ	JH	Jκ
DAB 494	48	IgM+ lymphoma	DJ/Myc	Amp.	Amp.	(split probe)	Rearr.
DAB 496	48	IgM+ lymphoma	DJ/Myc	Amp.	Amp.	(split probe)	Rearr.
DAB 473	68	IgM+ lymphoma	DJ/DJ	Amp.	Amp.	Rearr. and Amp.	Rearr.
DAB 538	69	IgM+ lymphoma	DJ/Myc	Amp.	Amp.	Rearr. and Amp.	Rearr.
DAB 361	70	IgM- lymphoma	DJ/Myc	Amp.	Amp.	Rearr. and Amp.	deleted
DAB 601	80	IgM+ lymphoma	DJ/Myc	Amp.	Amp.	Rearr. and Amp.	Rearr.
DAB 277	95	IgM+ lymphoma	DJ/Myc	Amp.	Amp.	Rearr. and Amp.	Rearr.

Amp. = Amplified signal
Rearr. = Rearrangement detected

Table 6. Control mice from DAB cohort. Age at the time of sacrifice is indicated in the second column. Cause of death (COD) is indicated in the third column. Genotype of each mouse is indicated in the last column.

Mouse ID	Age (days)	COD	Genotype	Mom	Dad
705#	95	T cell lymphoma	ATM-/- DJβ/Myc	76	115
685#	176	T cell lymphoma	ATM-/- DJβ/Myc	45	49
355#	265	T cell lymphoma	ATM-/- DJβ/Myc	76	115
235#	270	T cell lymphoma	ATM-/- DJβ/Myc	45	49
462	241	sacrificed	ATM+/- DJβ/Myc Eμ-Bcl-2	76	115
470	261	sacrificed	ATM+/- DJβ/Myc Eμ-Bcl-2	76	115
80	305	sacrificed	ATM+/- DJβ/Myc	bcl2 DJ/myc	ATM+/-
50	323	sacrificed	ATM+/+ Eμ-Bcl-2	bcl2 DJ/myc	ATM+/-
118	369	sacrificed	ATM+/- DJβ/Myc Eμ-Bcl-2	bcl2 DJ/myc	ATM+/-
115	375	sacrificed	ATM+/- DJβ/Myc Eμ-Bcl-2	bcl2 DJ/myc	ATM+/-
49	412	sacrificed	ATM+/- DJβ/Myc Eμ-Bcl-2	bcl2 DJ/myc	ATM+/-
76	414	sacrificed	ATM+/- DJβ/Myc	bcl2 DJ/myc	ATM+/-
43	417	sacrificed	ATM+/- DJβ/Myc Eμ-Bcl-2	bcl2 DJ/myc	ATM+/-
45	432	sacrificed	ATM+/- DJβ/Myc	bcl2 DJ/myc	ATM+/-

# indicates DA mice resulting from DAB cohort

2.3.3. ATM-Deficient Mature B Cell Lymphomas Arise From Naïve Mature B Cells Population

Analysis of surface marker by flow cytometry revealed that 13 of 16 B cell tumors obtained in DA, AB and DAB cohorts were B220+IgM+, suggesting they arose from mature B cells that did not undergo CSR. Of the tumors that were IgM+, 11 expressed Igκ and 2 (DAB473 and DAB496) expressed Igλ on their surface (Figure 12, Table 7). Three B cell lymphomas (AB36, AB225, and DAB361) lacked surface IgM and light chain expression (Table 7).

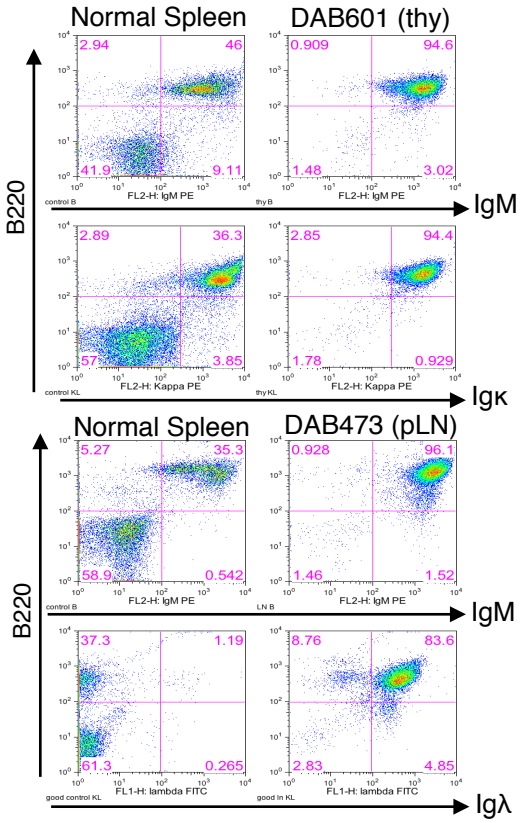


Figure 12. Example of flow cytometry analysis of tumor cells isolated from DAB601 thymus (thy) and DAB473 peripheral lymph nodes (pLN). Normal spleen was used as control. Antibodies used are indicated.

Table 7. Summary of cell surface marker analysis of tumor cells isolated from DA, AB, and DAB tumors by flow cytometry.

Mouse ID	Surface Marker
DA 64	B220+IgM+IgK+
DA 107	B220+IgM+IgK+
DA 360	B220+IgM+IgK+
DA 403	B220+IgM+IgK+
DA 607	B220+IgM+IgK+
AB 67	B220+IgM+
AB 36	B220-pan-Ig+
AB 143	B220+IgM+IgK+
AB 225	B220+IgM-IgL-
DAB 277	B220+IgM+IgK+
DAB 361	B220+IgM-IgL-
DAB 473	B220+IgM+Igλ+
DAB 494	B220+IgM+IgK+
DAB 496	B220+IgM+Igλ+
DAB 538	B220+IgM+IgK+
DAB 601	B220+IgM+IgK+

IgL- = no expression of IgL chain

To further verify B cell origin and clonality of these tumors, we examined the rearrangement status of the IgH and Igκ loci by Southern blot analysis, using probes that hybridize downstream of the JH region (JH 4-3 probe) and of the Jκ region (Jκ probe), respectively (Figure 13 and 14). The results showed that most tumors contained distinct rearranged bands for both JH and Jκ, indicative of monoclonal origin (Figure 13 and 14). Lack of rearranged JH bands in some tumors is likely due to aberrant processing of V(D)J joins in the absence of ATM, resulting in deletion of the region recognized by the probe (Lobrich and Jeggo, 2005). Amplification of the JH region was also detected, suggesting that the *IgH* locus may be involved in oncogenic translocations in these tumors as previously reported for C-NHEJ/p53-deficient pro-B cell tumors (Gao et al., 2000; Zhu et al., 2002) (Figure 13). Tumor AB36, which was surface IgM-negative, showed two Igκ rearrangements (Figure 15), and immunohystochemical analysis revealed it was

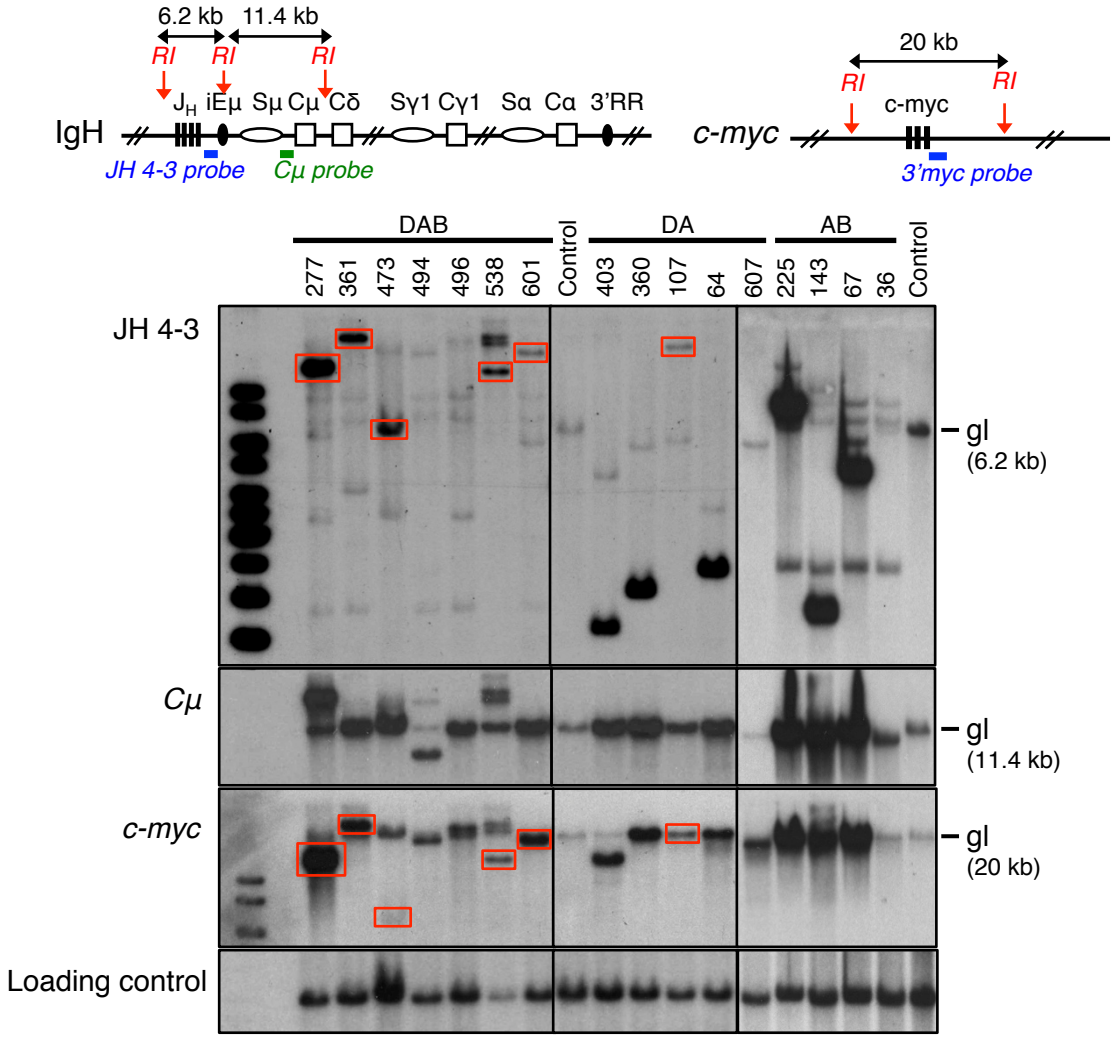


Figure 13. Southern blot analysis of tumor samples from DA, AB, and DAB cohorts. The membrane was hybridized with different probes as indicated on the left hand side. Top panel shows distinct rearranged bands and amplification of JH region. Second panel shows amplification of Cμ region. Third panel shows c-myc amplification and rearrangement. Red boxes indicate co-migrating c-myc and rearranged JH bands. Position and size of germline (gl) bands are indicated. Diagrams of the IgH and c-myc loci, with position of EcoRI (RI) sites and probes used are shown on top.

positive for intra-cytoplasmic Ig staining (Figure 15) suggesting that it may have arisen from a more mature, antibody-secreting cell type. Tumors DAB361 and AB225, which also lacked surface Ig expression, did not show Igκ rearrangements by Southern blot. However PCR analysis to assess the status of the lambda locus in these tumors revealed productive clonal Igλ

rearrangements (Figure 16). This indicates that these tumors did not derive from pro-B cells, which do not undergo IgL rearrangements, and that they likely deleted the Igκ locus during receptor editing (Wang et al., 2008).

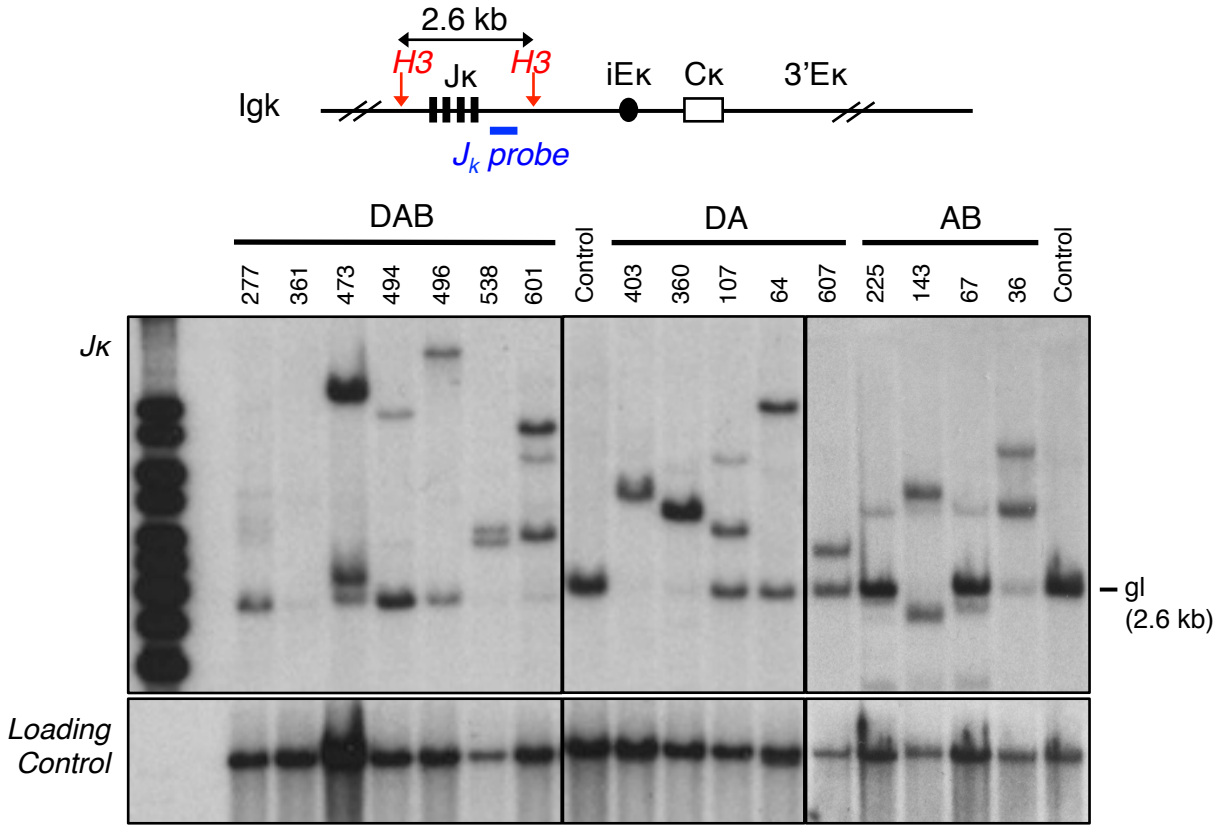


Figure 14. Southern blot analysis of tumor samples from all three cohorts. The membrane was hybridized with Jκ probe. Position and size of germline (gl) bands are indicated. Diagrams of the Jκ locus, with position of HindIII (H3) sites and probes used are shown on top.

To confirm that the tumors arose from cells that had not undergone CSR, the status of the Sμ switch region was analyzed by Southern blotting with a probe (Cμ probe) that recognizes a fragment between Sμ and Cμ, which is not altered by V(D)J recombination (Figure 13). This probe would detect a rearranged band if the cell of origin of the tumor had undergone CSR between Sμ and downstream S regions. All but two B cell lymphoma samples carried germ-line

unrearranged Cμ bands, suggesting that they derived from naïve mature B cells (Figure 13). We also performed sequencing of the JH intronic region and revealed that none of the B220+IgM+ tumors harbored mutations in this region, indicating that they had not undergone somatic hypermutation (Figure 17). Taken together, these findings suggest that ATM-deficient B cell tumors arose from naïve mature B cells in the periphery, which did not experience antigenic challenge.

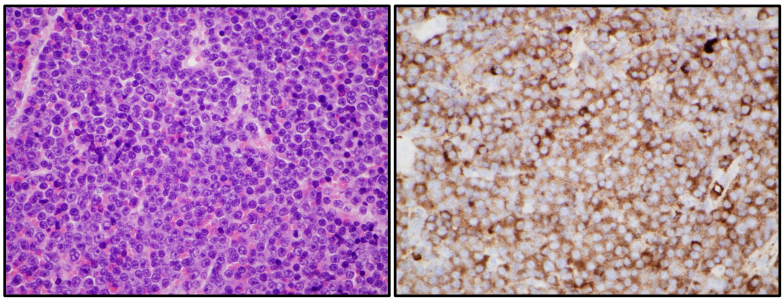


Figure 15. Immunohistochemistry analysis of AB36 gut mass tissue section. Left panel shows a tissue section stained with hematoxylin and eosin. Right panel shows a tissue section stained with anti-B220 antibodies.

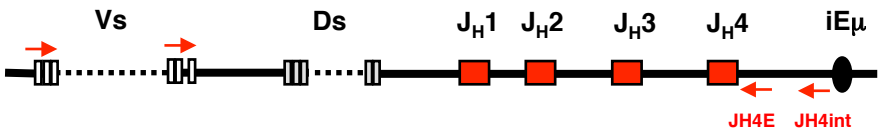
DAB361: productive V1-J3

CAGGCTGTTGTGACTCAGGAATCTGCACTCACCACATCACCTGGTGAAACAGTCACACTCACTTGT
CGCTCAAGTACTGGGGCTGTTACAACTAGTAACTATGCCAACTGGGTCCAAGAAAAACCAGATCATT
TATTCACTGGTCTAATAGGTGGTACCAACAACCGAGCTCCAGGTGTTCCTGCCAGATTCTCAGGCTC
CCTGATTGGAGACAAGGCCGCCCTCACCATCACAGGGGCACAGACTGAGGATGAGGCAATATATTT
CTGTGCTCTATGGTACAGCAACCATTTTATTTTCGGCAGTGGAACCAAGGTCACTGTCCTA

AB225: productive V1-J1

CAGGCTGTTGTGACTCAGGAATCTGCACTCACCACATCACCTGGTGAAACAGTCACACTCACTTGT
CGCTCAAGTACTGGGGCTGTTACAACTAGTAACTATGCCAACTGGGTCCAAGAAAAACCAGATCATT
TATTCACTGGTCTAATAGGTGGTACCAACAACCGAGCTCCAGGTGTTCCTGCCAGATTCTCAGGCTC
CCTGATTGGAGACAAGGCTGCCCTCACCATCACAGGGGCACAGACTGAGGATGAGGCAATATATTT
CTGTGCTCTATGGTACAGCAACCATTTGGTGTTCGGGGGAAGAACCAAACTGACTGTCC

Figure 16. Sequencing results of PCR products in DAB361 and AB225 tumor samples. The blue region represents Vλ genes. The green region represents Jλ genes. The red letters represent the junction associated with recombination.



TUMOR	FOR PRIMER	REV PRIMER	J SEGMENT	PRODUCT (bp)	SEQUENCE (bp)	MUT	STRAIN
DAB361	J558	JH4int	JH3	1900	1466	NO	B6
DAB494	J558	JH4E	JH2	1150	1102	NO	129
DAB494	7183	JH4E	JH2	1300	875	NO	129
DAB601	J558	JH4E	JH1	1400	1400	NO	129
AB67	J558	JH4int	JH4	1350	1132	NO	129
DA64	J558	JH4E	JH1	1400	1312	NO	129
DA107	J558	JH4int	JH1	1400	1000	NO	129
DA360	J558	JH4int	JH4	1350	1138	NO	129
DA403	J558	JH4int	JH4	1350	1113	NO	129
AB143	J606	JH4int	JH4	1350	1112	NO	129
DA607	7183	JH4E	JH2	1300	1075	NO	129

Figure 17. Somatic hypermutation analysis in a subset of tumors. Diagram of the V, D, and J gene segments and the positions of primers are shown in top panel. Results of the assay are shown in the bottom panel. Tumor samples are indicated in the first column. Forward and reverse primer that yielded PCR amplicons are indicated in second and third column, respectively. The utilization of J segment is shown in the forth column. Level of mutation (MUT) is indicated. Mouse strain alignment is indicated in the eighth column.

2.3.4. ATM-Deficient B Cell Lymphomas Harbor Complex Translocations Involving Chromosome 12 And 15.

One of the hallmark features of human and mouse B cell lymphomas is the clonal translocation juxtaposing oncogenes, such as *c-myc*, to *Ig* loci. To determine if this was also the case for B cell lymphomas in our mouse cohorts, Southern blot analysis was performed to examine the *c-myc* locus with a probe downstream of *c-myc* exon3 (3'Myc) (Figure 13). We detected *c-myc* amplification in all samples analyzed except DAB473 (see below) and AB36, which may have arisen from a different mechanism (Figure 13). In all but one DAB samples, the *c-myc* locus was also rearranged (Figure 13). To verify *c-myc* overexpression, we performed Northern blot analysis on a subset of tumor samples, which showed high levels of *c-myc* transcripts in all cases (Table 8-10).

Of note, Southern blot analysis showed that, in 6 of the tumors, the size of the *c-myc* fragment was the same as the size of the JH 4-3 fragment, suggesting that these tumors contain translocation that juxtaposes *c-myc* to *IgH*. To confirm this hypothesis, we performed fluorescent in situ hybridization (FISH) experiments on tumor metaphases. First, the metaphases were hybridized with paint probes specific for chromosome 12 and 15, where the *IgH* and *c-myc* loci reside, respectively. After image acquisition, the paint signals were stripped and the metaphases re-hybridized with BAC probes spanning *IgH* and *c-myc*. Translocation juxtaposing these two loci were present in all the tumors analyzed (Figure 18, Table 8-10).

Table 8. Detailed analysis of B cells lymphoma from DA cohort. Fold amplification (x amp.) of c-myc gene in indicated in the 9th row. The level of c-myc mRNA expression is indicated in the 10th row. Translocation breakpoint on chromosome 12 is indicated in the 12th row. Translocation breakpoint on chromosome 15 is indicated in the 13th row. The number of homology (hom) is indicated in samples with translocation junction derived from sequencing.

Mouse ID	DA64	DA107	DA360	DA403	DA607
Age (days)	95	103	96	135	90
Surface Marker	B220+ IgM+ IgK+	B220+ IgM+ IgK+	B220+ IgM+ IgK+	B220+ IgM+ IgK+	B220+ IgM+ IgK+
c-myc allele	DJ/DJ	DJ/WT	DJ/DJ	DJ/DJ	DJ/DJ
JH	Rearr. Amplified	Rearr. Not amp.	Rearr. Amplified	Rearr. Amplified	Rearr. Amplified
Cμ	GL Amplified	GL Amplified	GL Amplified	GL Amplified	GL Amplified
Iμ	GL Amplified	GL Amplified	GL Amplified	GL Amplified	GL Amplified
Jκ	Rearr.	Rearr.	Rearr.	Rearr.	Rearr.
c-myc	GL Amplified 6x amp.	GL Amplified 4.2x amp.	GL Amplified 6.7x amp.	Rearr. Amplified 3.8x amp.	GL Amplified 14x amp.
c-myc exp.	NA	2.6x	NA	11.8x amp.	NA
Cytogenetics	t(12;15) with 3'IgH and O21 (no D14) Chr15-based complicom 15;12;16 (AMP C10, AMP D14, AMP 3'IgH) 1 normal 12, 1 normal 15	t(12;15) with 3'IgH and O21 (no D14) Chr15-based complicom 15;12;? Chr15-based complicom ?;15;12 1 normal 12, 1 normal 15	t(12;15) with 3'IgH and O21 (no D14) Chr15-based complicom 15;12;? Or 15;1;16 Chr15-based complicom ?;1;15;16 1 normal 12, 1 normal 15	t(12;15;14) (AMP 3'IgH, AMP D14) 1 normal 12, 1 normal 15, chr15 with deletion DJb to 224kb DN Myc	t(12;8) no other complicon (AMP 3'IgH, AMP D14) 1 normal 12, 2 normal 15
Chr12 Bpt	214bp DN JH3	-	248bp DN JH2	-	-
Chr15 Bpt	114kb DN c-myc 2bp hom	est. 10kb DN c-myc	140kb DN c-myc no hom	est. >100kb DN c-myc	10-100kb DN c-myc
Translocation Allele	Not rearr. DJb	Aberrant rearr. DJb	Not rearr. DJb	Aberrant rearr. DJb	Aberrant rearr. DJb

Downstream is abbreviated as DN.
Germline is abbreviated as GL.
Rearranged allele is indicated as Rearr.

Table 9. Detailed analysis of B cells lymphoma from AB cohort. Fold amplification (x amp.) of c-myc gene in indicated in the 8th row. The level of c-myc mRNA expression is indicated in the 9th row. The estimated translocation breakpoint on chromosome 15 is indicated in the 11th row.

Mouse ID	AB67	AB36	AB143	AB225
Age (days)	104	297	116	84
Surface Marker	B220+ IgM+	B220- pan-Ig+	B220+ IgM+ IgK+	B220+ IgM- IgL-
JH	Rearr. Amplified	Deleted	Rearr. Amplified	Rearr. Amplified
Cμ	GL Amplified	GL Not Amplified	GL Amplified	GL Amplified
Iμ	GL Amplified	Deleted	GL Amplified	GL Amplified
Jκ	Rearr.	Rearr.	Rearr.	Deleted
c-myc	GL Amplified 9.3x amp.	GL Not Amplified	GL Amplified 8.5x amp.	GL Amplified 5.6x amp.
c-myc exp.	9x	NA	9.8x	NA
Chr12 Bpt	-	-	-	-
Chr15 Bpt	est. 100kb DN c-myc	-	est. >100kb DN c-myc	10-100kb DN c-myc
Cytogenetics	t(12;15) with 3'IgH and O21 Chr15-based complicom 15;12;16 1 normal 12, 1 normal 15	t(12;15) t(9;16) Robertsonian t(9;16;8) Robertsonian 1 normal 12, 1 normal 15	t(12;15) (AMP 3'IgH and D14) AMP 3'IgH and D14 on unknown chr. 1 normal 12, 2 normal 15	NA

Downstream is abbreviated as DN.
Germline is abbreviated as GL.
Rearranged allele is indicated as Rearr.

Table 10. Detailed analysis of B cells lymphoma from DAB cohort. Fold amplification (x amp.) of c-myc gene in indicated in the 9th row. The level of *c*-myc mRNA expression is indicated in the 10th row. Translocation breakpoint on chromosome 12 is indicated in the 11th row. Translocation breakpoint on chromosome 15 is indicated in the 12th row. The number of homology (hom) is indicated in samples with translocation junction derived from sequencing.

Mouse ID	DAB277	DAB361	DAB473	DAB494	DAB496	DAB538	DAB601
Age (days)	95	70	68	48	48	69	80
Surface Marker	B220+ IgM+ IgK+	B220+ IgM- IgL-	B220+ IgM+ Igλ+	B220+ IgM+ IgK+	B220+ IgM+ Igλ+	B220+ IgM+ IgK+	B220+ IgM+ IgK+
c-myc allele	DJ/WT	DJ/WT	DJ/DJ	DJ/WT	DJ/WT	DJ/WT	DJ/WT
JH	Rearr. Amplified	Rearr. Amplified	Rearr. Amplified	deleted	deleted	Rearr. Amplified	Rearr. Amplified
Cμ	Rearr. Amplified	GL Amplified	GL Amplified	Rearr. Amplified	GL Amplified	GL Amplified	GL Amplified
Iμ	Rearr. Amplified	GL Amplified	GL Amplified	deleted	GL Amplified	GL Amplified	GL Amplified
Jκ	Rearr.	Rearr.	Rearr.	Rearr.	Rearr.	Rearr.	Rearr.
c-myc	Rearr. 18x amp.	Rearr. 7.5x amp.	Rearr. 2x amp.	Rearr. 4.8x amp.	Rearr. 4.5x amp.	Rearr. 4.3x amp.	Rearr. 6x amp.
c-myc exp.	NA	NA	4.3x	7.1x	NA	NA	8.7x
Chr12 Bpt	-	27bp DN JH4	507bp DN JH4	860bp DN JH4	890bp DN JH4	JH3	780bp DN JH4
Chr15 Bpt	est. between Myc A-D	5.5kb DN of c-myc 2bp hom	99bp DN of c-myc 2bp hom	4kb DN of c-myc 3bp hom	2.8kb DN of c-myc 3bp hom	1.1kb DN of c-myc no hom	4kb DN of c-myc 2bp hom
Cytogenetics	Chr12-based complicom 1 normal 12, 2 normal 15	Chr12-based complicom 1 normal 12, 2 normal 15	No metaphase	Chr15-based complicom t(12;15) with 3'IgH and L21 1 normal 12, 1 normal 15	Chr12-based complicom 1 normal 12, 2 normal 15	No metaphase	Chr12-based complicom 1 normal 12, 2 normal 15
Translocation Allele	inconclusive	Aberrant Rearr. DJb	Aberrant Rearr. DJb	Aberrant Rearr. DJb	Aberrant Rearr. DJb	Aberrant Rearr. DJb	Aberrant Rearr. DJb

Downstream is abbreviated as DN.
Germline is abbreviated as GL.
Rearranged allele is indicated as Rearr.

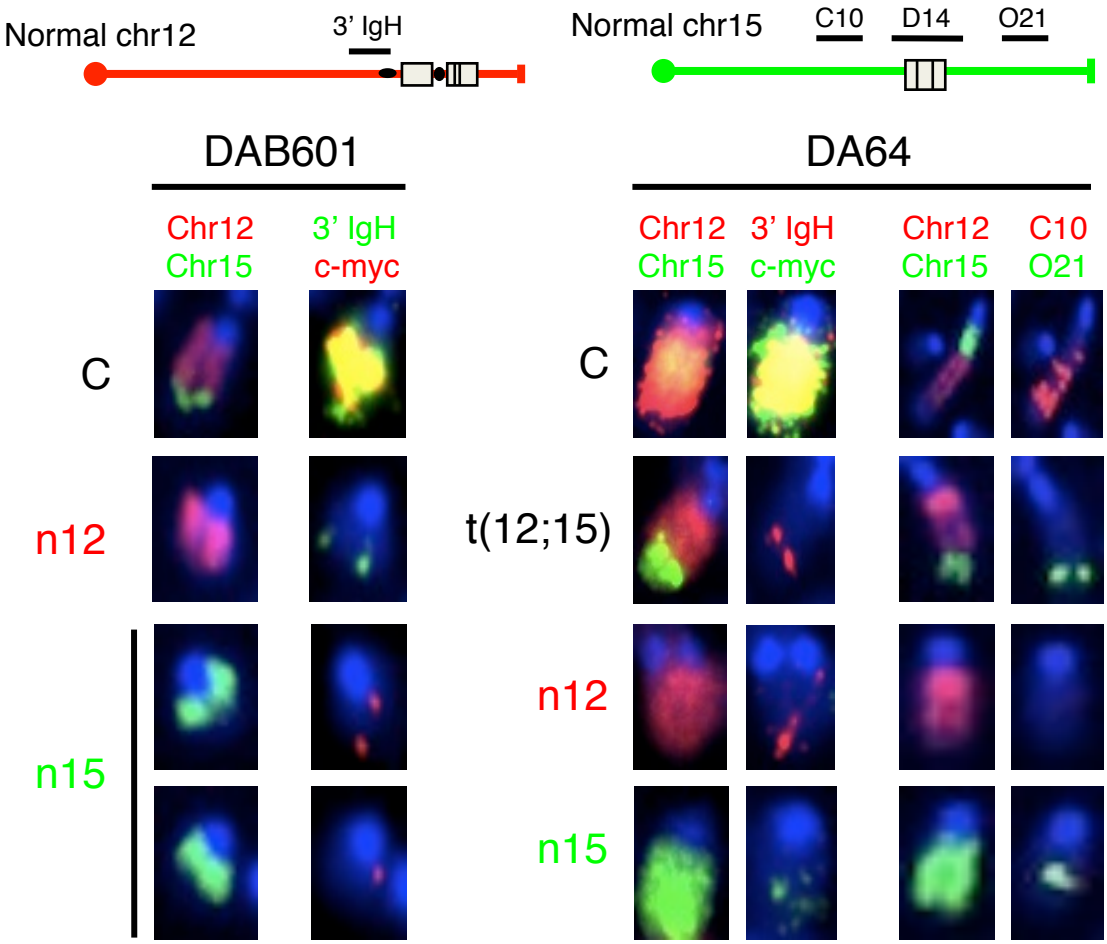


Figure 18. FISH analysis on metaphases from DAB601 and DA64 tumor samples. Details of 12;15 complicons (C), 12;15 translocation with no amplification t(12;15), and normal chromosome 12 (n12) and 15 (n15) are shown. Probes used are indicated. The diagrams on top represent normal chromosome (chr) 12 and 15, with location of the BAC probes.

In line with Southern blot results, complex translocations with gene amplification (complicons) were detected in all tumors analyzed, similarly to what previously described for pro-B cell tumors developing in C-NHEJ/p53-deficient mice (Difilippantonio et al., 2002; Zhu et al., 2002). We identified two types of complicons, that we refer to as 15-based (retain the centromeric portion of chromosome 15) and 12-based (retain the centromeric portion of chromosome 12) complicons (Figure 18). Both types of complicons harbor amplification of *IgH* and *c-myc* and were present in all three cohorts. In tumors where a 15-based complicon was present, a T(12;15) translocation juxtaposing the 3' region of *IgH* to chromosome 15 regions downstream of *c-myc* was also detected, in addition to one normal chromosome 12 and one normal chromosome 15. The T(12;15) translocation was not detected in tumors with 12-based complicon, which instead carried one normal chromosome 12 and two normal chromosome 15. These two types of complicon were previously reported in pro-B cell lymphomas (Difilippantonio et al., 2002; Zhu et al., 2002), suggesting that both pro-B and mature B cell lymphomas may undergo similar transformation events that lead to lymphoma development.

2.3.5. Translocation Breakpoints Downstream Of *c-myc* Lead To Gene Amplification Via Breakage-Fusion-Bridge Mechanism

The *c-myc* gene amplification in C-NHEJ/p53-deficient pro B cell lymphomas is proposed to arise via a breakage-fusion-bridge mechanism that require the formation of dicentrics resulting from the fusion of RAG-initiated *IgH* DSBs to sequences downstream of *c-myc* (Zhu et al., 2002). To test if this was also the case in ATM-deficient B cell lymphomas, detailed analysis of the *c-myc* locus with probes upstream of *c-myc* exon 1 (MycA), in *c-myc* intron 1 (MycD) and downstream of *c-myc* exon 3 (3'Myc) was performed. In some tumors (DA64, DA107, DA360, AB67, AB143, and AB225), lack of rearrangement yet amplification of *c-myc* was detected with all three probes, suggesting the translocation junctions lay outside of the 20kb EcoRI fragment (Figure 19). Other tumors (DAB277, DAB361, DAB494, DAB496, DAB538, and DAB601) showed both rearrangement and *c-myc* amplification with MycD and/or 3'Myc probes, suggesting that breakpoints were within the *c-myc* locus (Figure 19).

To confirm that the translocation junctions lay downstream of the *c-myc* locus, we employed comparative genomic hybridization (CGH) to examine the level of amplification at *c-myc* and *IgH* loci in a few samples from each cohort. This analysis showed that amplification in the *IgH* locus began downstream of D gene segments in all samples analyzed (DAB494, DAB496, DAB601, and DA64) whereas amplification on chromosome 15 started either not far downstream of *c-myc* (DAB494, DAB496, DAB601) or in the *Pvt1* region, 100kb downstream of *c-myc* (DA64), according to the results of the Southern blot data (**Figure 20**). These data helped us define the region where translocation breakpoints were located. Translocation junctions from these and four more tumors were then cloned by PCR approach and sequenced.

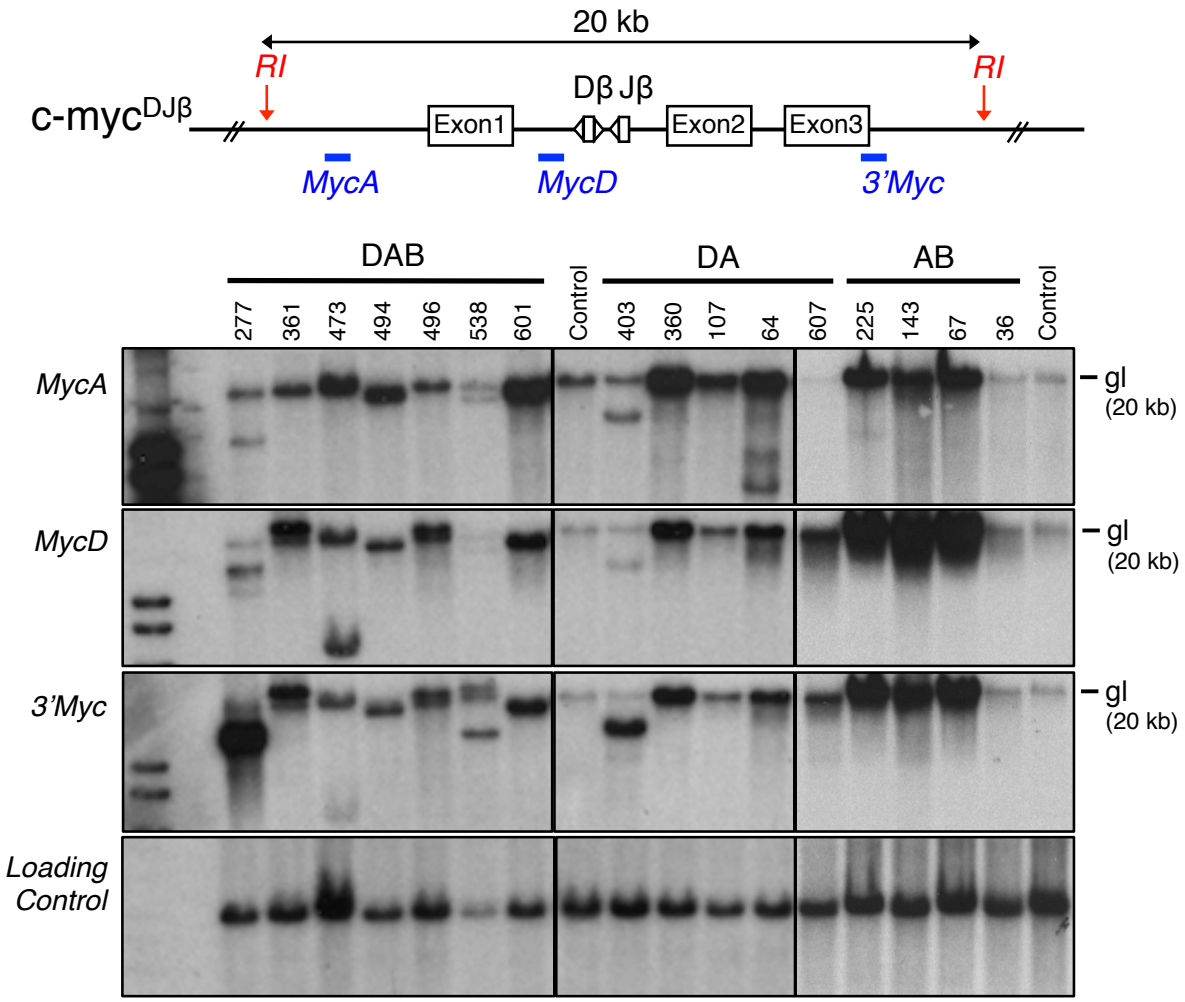


Figure 19. Detailed Southern blot analysis of the *c-myc* locus. The membrane was sequentially hybridized with three probes spanning c-myc represented in the diagram on top and indicated on the left hand side. Position and size of germline (gl) bands are indicated. Note that the top bands recognized by 3'Myc probe in tumor DAB538 results from incomplete digestion of genomic DNA as repeat Southern blots do not show such pattern.

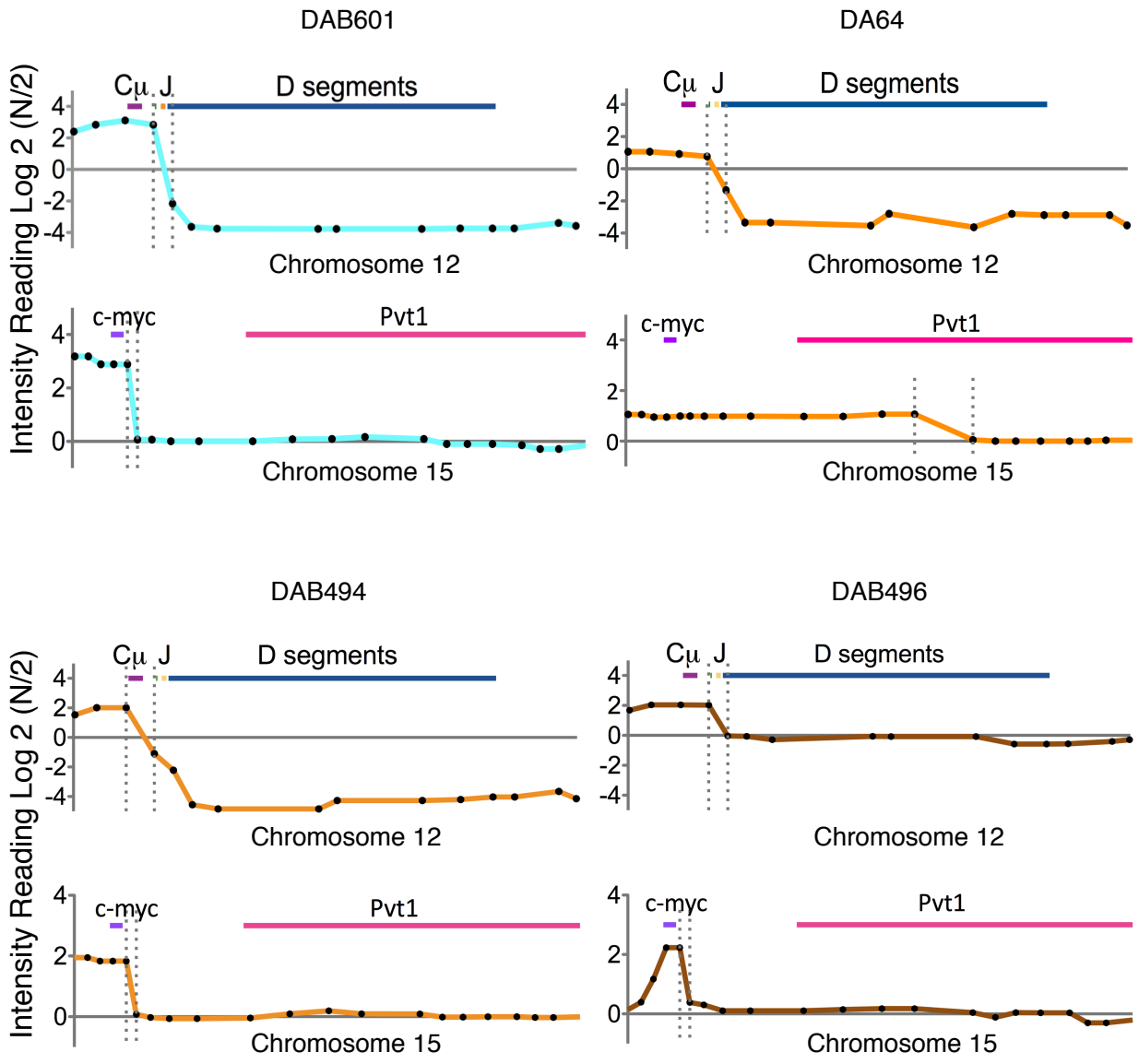


Figure 20. CGH analysis on a subset of tumor samples. Dotted lines represent the transition point where amplification begins. Each circle represents individual CGH probe.

The sequencing results showed juxtaposition of chromosome 12 sequences between Eμ and JH segments to chromosome 15 sequences downstream of *c-myc,* ranging from 99bp to 140kb **(**Figure 21, Table 8-10). The orientation of all sequenced translocation junctions was consistent with amplification of both loci via a breakage-fusion-bridge mechanism. These sequencing results also corresponded with rearrangement and amplification patterns detected in Southern blot analysis. For example, the lack of amplification in the rearranged *c-myc* fragment detected by Southern blot with 3'Myc probe in DAB473 is due to the translocation breakpoint being immediately downstream of *c-myc* exon 3, therefore eliminating most of the genomic region recognizable by the probe **(**Figure 21, Table 8-10). Likewise, the lack of amplified and rearranged fragments detectable by the JH 4-3 probe in tumors DAB494 and DAB496 is the result of translocation junctions on chromosome 12 deleting most of the region recognized by the probe **(**Figure 21, Table 8-10). In tumors DAB361, DAB473, and DAB538, the lack of rearranged band detected by MycA probe suggests a secondary rearrangement upstream of *c-myc,* in addition to the oncogenic translocation junctions in the *IgH* locus.

It is notable that the five translocation junctions lie very close to *c-myc,* a region not previously identified as translocation partners in C-NHEJ/p53-deficient pro-B tumors (Figure 20, label 1-5). Study of the precise location of these junctions indicated that all were within a few hundred basepairs from strong cryptic RSSs, with the closest one overlapping with a cryptic RSS with an addition of one nontemplated nucleotide (Table 11). The distance between the translocation junction and the cryptic RSS is compatible with the resection observed in the absence of H2AX and 53BP1; both are ATM substrates (Helmink and Sleckman, 2012), and presumably would be the case in ATM deficiency. Similarly, the two translocation junctions

located far downstream of *c-myc* in *Pvt1* also mapped in the vicinity of cryptic-RSSs although the distance between the translocation junctions and cryptic-RSSs was further than those located close to *c-myc* (Table 11).

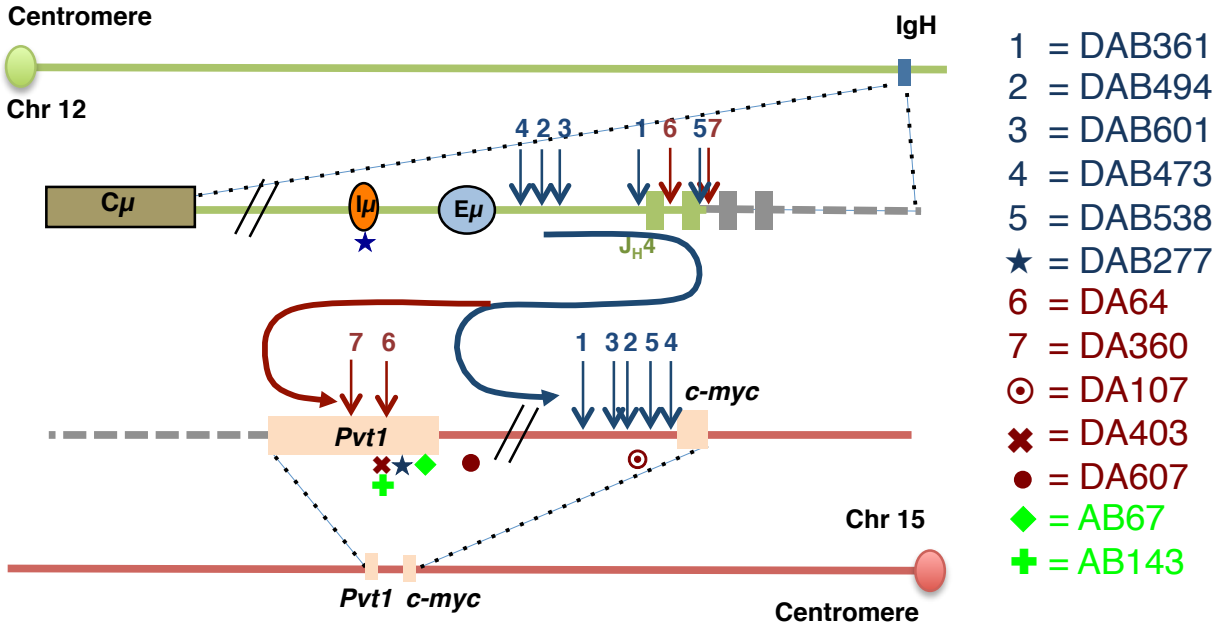


Figure 21. Schematic of translocation breakpoints in tumor samples. Data are either from direct sequencing of junctions (shown as numbers) or estimated (shown as symbols) from Southern blot analysis. See text for details.

To estimate translocation breakpoint from the remaining tumors we used Southern blot analysis with various digests and probes. SacI digestion of tumor DA107 DNA showed a rearranged band recognized by 3'Myc probe, suggesting the translocation lay within this 10kb region downstream of *c-myc* (Figure 21, Figure 22). Examination of the *Pvt1* locus, which from our data and previous reports (Zhu et al., 2002) is frequently involved in complicon formation with *IgH*, allowed us to estimate that translocation breakpoints for DA403, AB67, AB143, and DAB277 to a region at least 100kb downstream of c-myc (Figure 23). For DA607, we determined

that the translocation breakpoint located in the 97kb region downstream of *c-myc* (Figure 21, Figure 23).

Table 11. **Mapping of translocation junctions from a subset of DA and DAB tumors in relation to cryptic RSSs near translocation breakpoint.** Coordinates (mm9) of translocation breakpoint downstream of *c-myc* on chromosome 15 is indicated in the 2nd column. RIC scores indicate the functionality of cryptic RSSs, which ranges from -1000 (very bad) to 0 (very good) (Cowell et al., 2002). The distance (in basepair units) between cryptic RSS cut site and the observed translocation breakpoint is indicated in the last column.

Tumor	Translocation Junctions		Closest Cryptic RSS					
	coordinate	Strand	Type	RIC Score	Start	End	Strand	Distance
DAB473	61,822,015	-	23 RSS	-56.26	61,822,577	61,822,615	-	600
DAB361	61,827,453	-	12 RSS	-34.49	61,827,543	61,827,570	-	117
DAB494	61,825,906	-	23 RSS	-54.68	61,825,932	61,825,970	-	64
DAB601	61,825,971	-	23 RSS	-54.68	61,825,932	61,825,970	-	1*
DAB538	61,823,031	-	23 RSS	-55.64	61,823,369	61,823,407	+	338
DA64	61,936,069	-	23 RSS	-51.05	61,936,660	61,936,698	+	591
DA360	61,962,080	-	23 RSS	-55.07	61,963,067	61,963,105	-	1,025

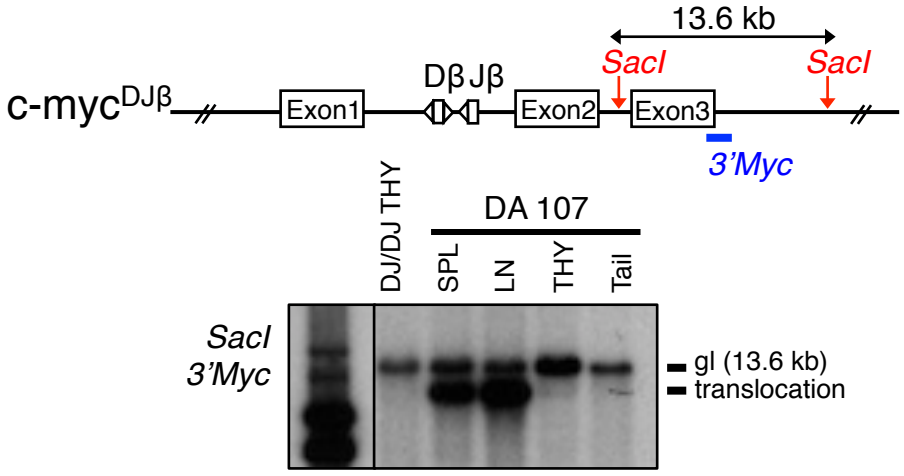


Figure 22. Southern blot analysis of DA107 showing translocation close to the *c-myc* locus. Position and size of germline (gl) band are indicted. Diagram of the c-myc locus with position of SacI sites and the probe used is shown on top.

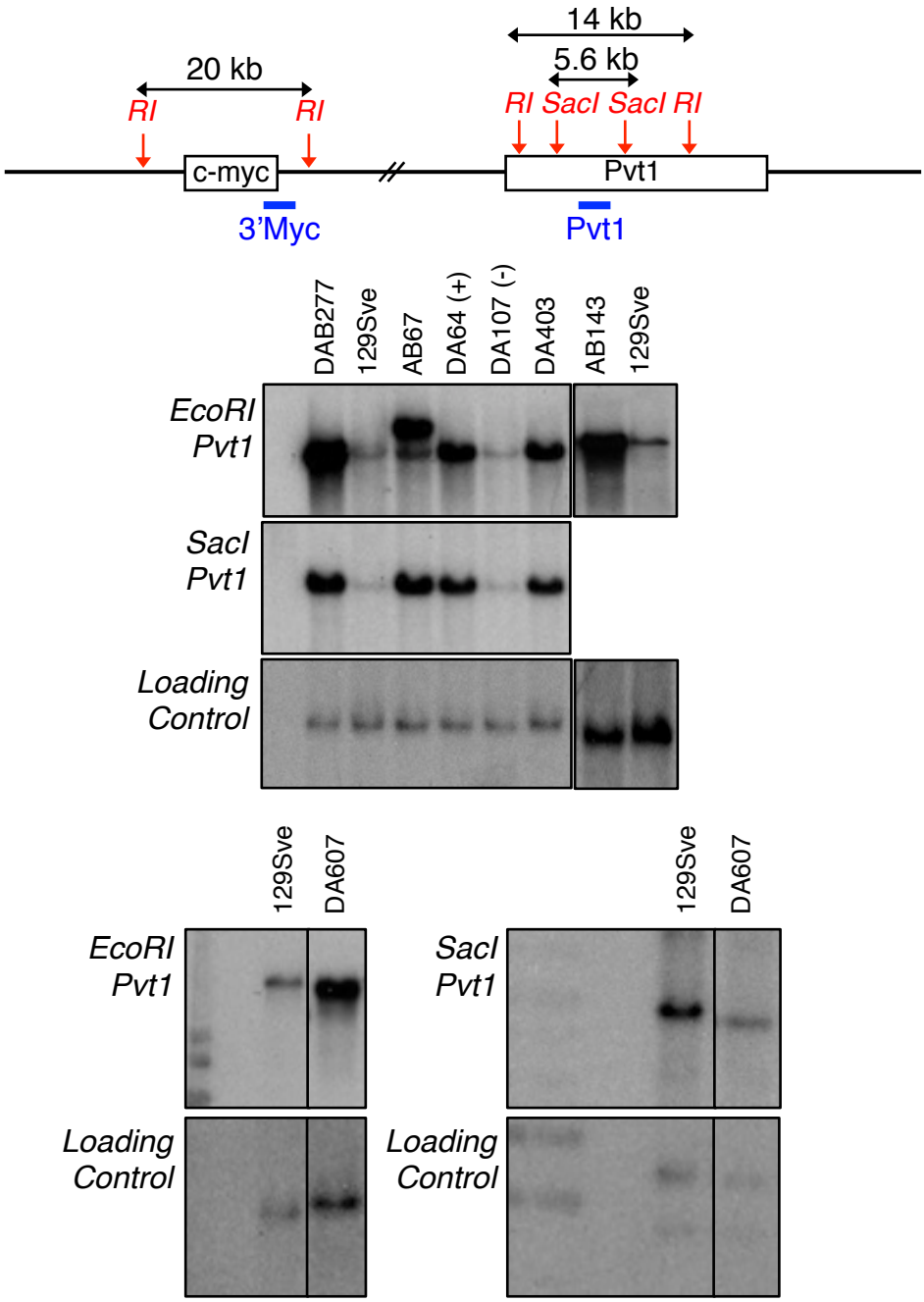


Figure 23. Southern blot of genomic DNA from DA and AB tumors showing amplification in *Pvt1* in AB143, DAB277, DA403, and DA607 samples. AB67 showed amplification of germline *Pvt1* fragment with SacI digestion but a rearranged and amplified EcoRI band with the same probe. DNA from WT mice (129Sve) was used as a control for normal *Pvt1* level. DNA from DA64 and DA107 tumors was used as a positive control (+) or negative control (-) for amplification, respectively. Diagram of *c-myc* and *Pvt1* locus showing EcoRI and SacI sites and probes used are shown on top. Digestion and probe used are indicated on the left hand side.

2.3.6. Translocation In DA And DAB B Cell Lymphoma Do Not Involve RAG Target Sequence

Sequencing of translocation breakpoints and Southern blot analysis suggested that the RAG target sequence inserted in intron 1 of *c-myc* was not the main driver of oncogenic translocations in DA and DAB tumors. Given that six out of seven DAB mice and mouse DA107 were heterozygous for the *c-myc*$^{DJ\beta}$ allele, we sought to determine which *c-myc* allele was involved in translocations in these tumors. To this end, Southern blot analysis was performed on genomic DNA digested with SacI using the MycD probe, which hybridizes in *c-myc* intron 1 upstream of the inserted DJβ RAG target sequence (Figure 24). This strategy allows us to distinguish between wild-type *c-myc* allele, unrearranged *c-myc*$^{DJ\beta}$ allele, and rearranged *c-myc*$^{DJ\beta}$ allele. Based on our previous characterization of these tumors, we are able to link amplification of the region recognized by MycD probe with the allele harboring the translocation.

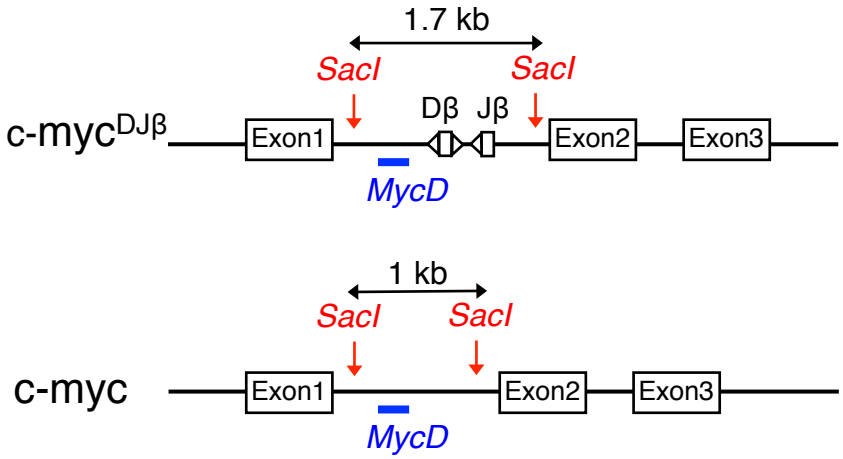


Figure 24. Map of *c-myc* locus showing Southern blot strategy with SacI sites and location of the MycD probe.

The results showed that in all tumors for which we could get unequivocal results (6 out of 7) the translocation involved the *c-myc*$^{DJ\beta}$ allele, which was invariably aberrantly rearranged,

in line with Southern blot data indicating secondary breakpoints between MycA and MycD or 3'Myc probes. With the same approach we were also able to analyze the status of the DJb cassette on the allele involved in translocation in the six *c-myc*$^{DJ\beta}$ homozygous tumors (five out of six DA, see table 1, and DAB473): two of them harbored translocation on unrearranged *c-myc*$^{DJ\beta}$ alleles (DA64 and DA360) whereas translocations in the other four tumors were associated with aberrant rearrangements of the genomic region containing the DJβ sequence (Figure 25). We conclude that even though the translocation does not directly involve the DJβ cassette, the presence of this sequence within c-myc renders the locus more susceptible to translocation.

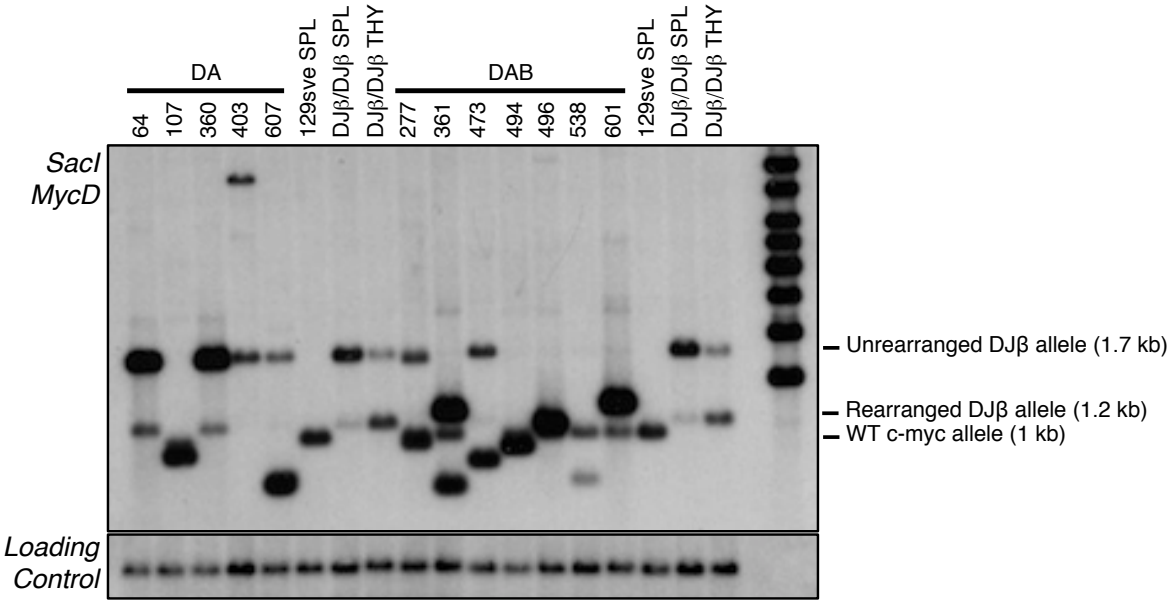


Figure 25. Southern blot of genomic DNA from DA and DAB tumors. DNA from spleen (SPL) or thymus (THY) of WT mice (129Sve) and mice homozygous for c-myc$^{DJ\beta}$ allele (DJβ/DJβ) were used as controls. Running position of WT c-myc, unrearranged c-myc^{DJB} and rearranged c-myc^{DJB} are indicated.

2.4. Discussion

2.4.1. Mouse Models That Spontaneously Develop IgM$^+$ B Cell Lymphomas In The Absence Of ATM.

ATM-deficient humans are predisposed to both T and B cell lymphomas but ATM-deficient mice develop only thymic lymphomas (Ambrose and Gatti, 2013). Studies of recurrent translocations in ATM-deficient thymic lymphomas revealed that they involved the TCRδ locus, a finding quite unexpected at the time as it was generally thought that such translocations would involve the TCRα locus (Zha et al., 2010). Subsequently, TCRδ locus translocations were found to be the most common RAG-initiated oncogenic translocations in human T-cell acute lymphoblastic leukemias (T-ALLs), which is also the most frequent tumor type in pediatric AT patients (Larmonie et al., 2013; Taylor et al., 1996). To gain potential insights into the roles of ATM in suppressing B cell lymphoma, we generated three different mouse models that develop similar peripheral B cell lymphomas in the context of ATM deficiency. The DA and AB model are based, respectively, on the introduction of either a *Tcr* DJβ V(D)J recombination cassette into *c-myc* intron 1 (Ranganath et al., 2008) or an E*μ-bcl-2* transgene (Strasser et al., 1994) into the ATM-deficient background (Borghesani et al., 2000). B cell lymphomas arise, respectively, in approximately 25% of the DA and 50% of the AB mice. The DAB model genetically combines the DA and AB models and develops peripheral B cell lymphomas rapidly and with 100% penetrance. Nearly all of the spontaneous clonal B cell malignancies that arose in each of these models were IgM$^+$ and harbored clonal translocations from the *IgH* JH region to sequences downstream of *c-myc* to form dicentric translocations that ultimately led to the *IgH/c-myc* amplification in the context of complicons.

2.4.2. Recurrent Oncogenic Translocations And Amplifications In ATM-Deficient B Cell Lymphomas.

It is striking that nearly all DA, AB, and DAB peripheral B cells lymphomas were IgM^+ and shared a common oncogenic translocation that joined sequences within or very near the *IgH* JH region to sequences downstream of *c-myc* to create dicentrics with chromosome 12-based or chromosome 15-based complicons. These translocations provide the basis for *c-myc* gene amplifications via the breakage-fusion-bridge-based mechanism. The general structure of the *IgH/c-myc* complicons and the presence or absence of other associated translocations (e.g. presence of a T(12;15) in tumors with a chromosome 15-based complicon but not those with a chromosome 12-based complicon) in these tumors were strikingly similar to the types and patterns of chromosomal aberrations found in RAG-dependent pro-B lymphomas that arise in C-NHEJ/p53-deficient mice (Zhu et al., 2002; Difilippantonio et al., 2002). In this regard, the clustering of the *IgH* translocation breakpoints in and around a small JH region in both the ATM-deficient and the C-NHEJ/p53-deficient background indicates that the *IgH* breaks that initiate formation of the complicons of ATM-deficient IgM^+ B cell lymphomas are RAG-initiated.

The question of how RAG-initiated chromosomal aberrations generated in pro-B cells leads to oncogenic translocations recurrently found in certain types of human peripheral B cell lymphomas has long been discussed. Based on striking similarities between the complicons in ATM-deficient mouse peripheral B cell lymphomas and those of C-NHEJ/p53-deficient mouse pro-B tumors, we suggest that the generation of the *IgH/c-myc* complicons is initiated in ATM-deficient progenitor B cells. How then do we explain why C-NHEJ/p53-deficient mice succumb to pro-B lymphomas but the susceptible ATM-deficient mice succumb to peripheral B cell

lymphomas? In this regard, C-NHEJ-deficiency abrogates B cell development at the pro-B stage due to a complete block in V(D)J recombination (Gostissa et al., 2011). Thus, in the absence of both C-NHEJ and p53, pro-B cells replicate persistent RAG-initiated DSBs due to the G1 checkpoint defect, cycling them into dicentrics, complicons, and ultimately leading to pro-B cell lymphomas (Zhu et al., 2002). Like dual C-NHEJ/p53-defciency, ATM-deficiency alone results in both V(D)J recombination joining defects and impaired p53 activation that causes G1 checkpoint defects (Helmink and Sleckman, 2012). However, unlike C-NHEJ deficiency, ATM deficiency does not completely block V(D)J joining (Barlow et al., 1996; Elson et al., 1996; Xu et al., 1996). Thus, we propose that some ATM-deficient pro-B cells generate dicentric chromosomes from replication of persistent RAG-initiated DSBs on one IgH allele but that productive V(D)J rearrangements on the other allele allows B cell harboring dicentrics, and potentially their downstream translocation/amplification products, to survive to the IgM^+ B cell stage where they contribute to B cell lymphomagenesis.

We hypothesized that B cell lymphomas in ATM-deficient background can be promoted by increased DSBs level in oncogenes such as *c-myc* or enhanced survival of ATM-deficient B cells. The AB B cell lymphoma model was based on the hypothesis that an Eμ-bcl2 transgene could promote survival of ATM-deficient developing or mature B cells that are eliminated via apoptosis from contributing to B cell lymphomas. Based on the success of the AB model, this general hypothesis appears validated. On the other hand, the manner in which the DJβ cassette targeted into *c-myc* promotes peripheral B cell lymphomas in DA or DAB ATM-deficient backgrounds is unlikely to result from the originally envisioned mechanism of increasing DSBs and translocations within intron one of *c-myc*. We found that all of the primary dicentric

junctions on chromosome 15 that gave rise to the *IgH/c-myc* complicons occurred downstream (telomeric) of the *c-myc* gene and did not involve the DJβ cassette. In this regard, when the c-myc^{DJβ} mice were generated, the DJβ cassette was targeted into *c-myc* intron one to provide both transcription to target RAG and also to provide a potential break site in the first intron, a region where translocation junctions to *IgH* often occur in human B cell lymphomas (Wilda et al., 2004). However, our studies of a mouse model that develops B cell lymphomas in association with recurrent *IgH/c-myc* translocations revealed that all translocation junctions are mapped to a region upstream of *c-myc* promoter and not into intron one (Gostissa et al., 2009). Currently, it seems likely that translocations into mouse *c-myc* intron one may inactivate the *c-myc* gene by disrupting an alternative promoter located upstream of *c-myc* exon 2 (Bernard et al., 1983; Saito et al., 1983).

While we have ruled out direct participation of the DJβ cassette within *c-myc* in translocation junctions, our findings still implicate the cassette as a contributing factor to the process. Thus, all (6/6) *IgH* translocation junctions that could unequivocally be analyzed in DA or DAB mice that were heterozygous for the c-myc^{DJβ} allele involved sequences downstream of *c-myc* on the DJβ cassette-targeted allele, a highly significant finding (*p* value = 0.0143). In this context, it is further notable that a substantial fraction of the DAB and DA junctions clustered within a 4kb region just downstream of *c-myc,* a region not identified as having complicon junctions in C-NHEJ/p53-deficient pro-B tumors in which junctions occurred much further downstream (Zhu et al., 2002). Examination of the precise location of junctions within the hotspot downstream region just downstream of *c-myc* indicated that all were within a few hundred bp of a set of strong cryptic RSSs, with one actually joined within 1bp of a cryptic RSS.

Based on these observations, we hypothesize that DJβ cassette indirectly contributes to downstream DSBs by providing proximal 12/23 RSSs for pairing with closely linked cryptic downstream RSSs and, thereby, enhances RAG cutting at these sites (and potentially also at positions further downstream). In the absence of ATM, such RAG-initiated DSBs might be aberrantly repaired as previously described (Huang et al., 2007; Bredemeyer et al., 2006). In this regard, B cell lymphomas would result only from B lymphocytes in which DJβ DSBs are resolved within the cassette (for example, by open and shut joins), thereby preserving *c-myc* integrity on that allele while partner DSBs at downstream cryptic RSS are not rejoined and could lead to chromosome breaks and dicentrics. Similar mechanisms may also lead to the generation of "off-target" RAG-initiated DSBs that contribute to T and B cell lymphomas in humans.

CHAPTER 3: MECHANISMS OF CHROMOSOMAL TRANSLOCATION IN ATM-DEFICIENT PERIPHERAL B CELLS

3.1. Results

3.1.1. ATM Suppresses Genome-Wide Translocations In Stimulated B Cells

Findings from the tumor cohorts prompted us to investigate the role of ATM in suppressing translocations in mature B cells. To this end, we employed high-throughput genome-wide translocation sequencing (HTGTS) to map translocation junctions in *in vitro* stimulated splenic B cells (Chiarle et al., 2011). B cells were isolated from wild-type (WT) and ATM-deficient mice harboring 25 consecutive I-SceI recognition sites (referred to as 25xI-SceI cassette) inserted into *c-myc* intron 1, in the same location as the RAG target sequence used for the tumor studies. Upon expression of I-SceI enzyme in B cells, specific DSBs are efficiently generated at these sites, and can be used as anchor to capture genome-wide translocations (Chiarle et al. 2011). We used primers located about 100 bp centromeric to the 25xI-SceI cassette to detect translocation junctions involving broken ends (BEs) on the 5' side of c-myc I-SceI DSBs (''5' BEs''). Based on convention, translocation partner sequences joined to 5' BEs are in (+) orientation if read from the junction in centromere to telomere direction and in (-) orientation if read in the opposite direction (Chiarle et al., 2011). HTGTS libraries were generated from WT and $ATM^{-/-}$ B cells, stimulated for CSR with α-CD40 and IL4. Results from 3 independent samples per each genotype yielded comparable numbers of total junctions, which were pooled together for analysis (Figure 6A, 6B). In our previous studies (Chiarle et al., 2011), we found that most of HTGTS reads from WT c-$myc^{25xI\text{-}SceI}$ cultured B cells mapped to a 2Mb region spanning the I-SceI breaksite on chromosome 15, with the vast majority of them corresponding to resection events of a single I-SceI DSB rather than to translocations joining two independent DSBs. Here, in the absence of ATM, we noticed an accumulation of junctions on the breaksite chromosome that was even more striking than in WT cells, and spanned the

whole length of the chromosome (Figure 26), similarly to what previously reported in IR-treated, G1-arrested ATM-deficient pro-B cell lines (Zhang et al., 2012). Moreover, only 64% of junctions was in the 2Mb region around the breaksite in ATM$^{-/-}$ cells, as compared to 78% in WT cells, and correspondingly a larger percentage (36% vs 22%) of junctions mapped elsewhere in the genome (Figure 27). We then determined how translocations distributed across the genome, and found that the number of 2 Mb bins containing 1, 2, 3, or 4 junctions, which may represent background level, were similar in both WT and ATM-deficient cells. However, ATM-deficient cells had a lower number of bins containing no translocation junction as well as a higher number of bins containing more than 4 translocation junctions (**Figure 28**), which may represent potential hotspots. These results indicate that ATM suppresses genome-wide translocations to maintain genomic stability in B cells.

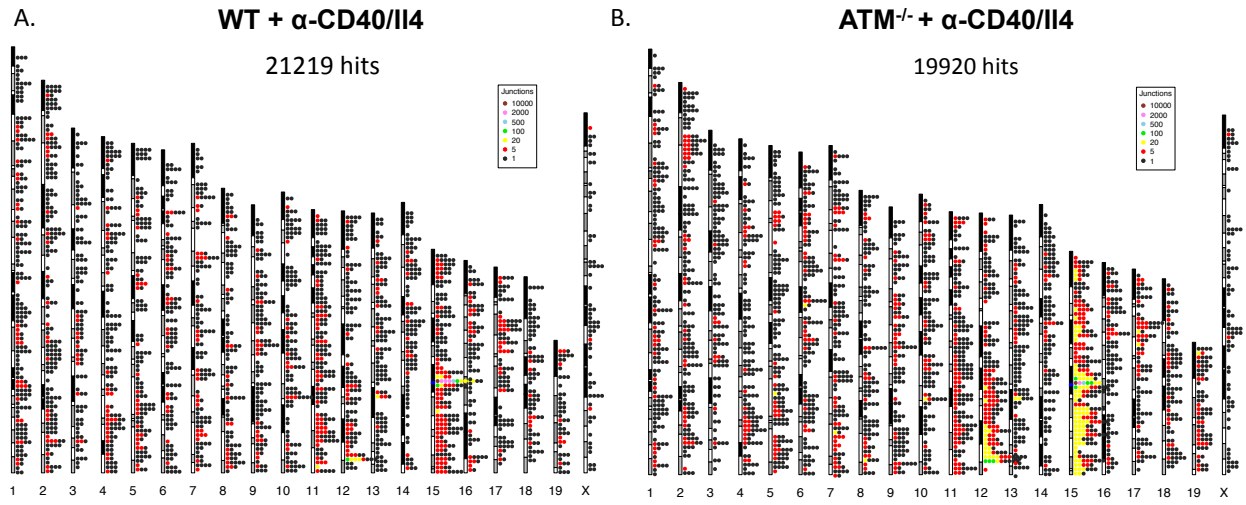


Figure 26. Distributions of genome-wide translocation in HTGTS libraries from α-CD40/IL4-stimulated (A) WT and (B) ATM-deficient B cells. Each plot represents a pool from three independent experiments. Genotype, treatment, and total number of junctions in each library are shown above each plot. The bin size is 2 Mb. Clusters of translocation hits are represented by color dots as showed in the square.

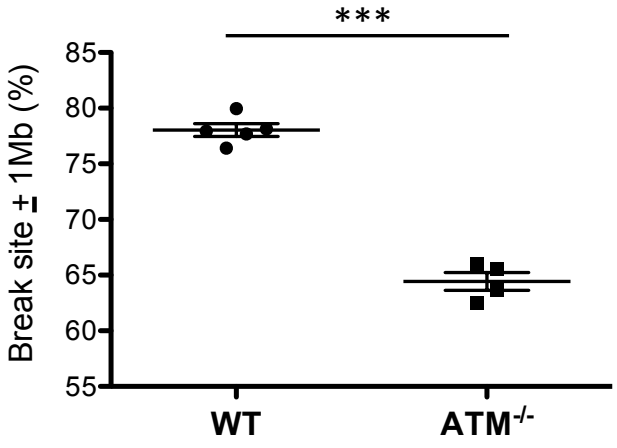


Figure 27. Comparison of numbers of translocation junctions mapped to a ±1Mb region around I-SceI *c-myc* breaksite. Statistical analysis was done with independent repeats from the pooled libraries showed in Figure 26.

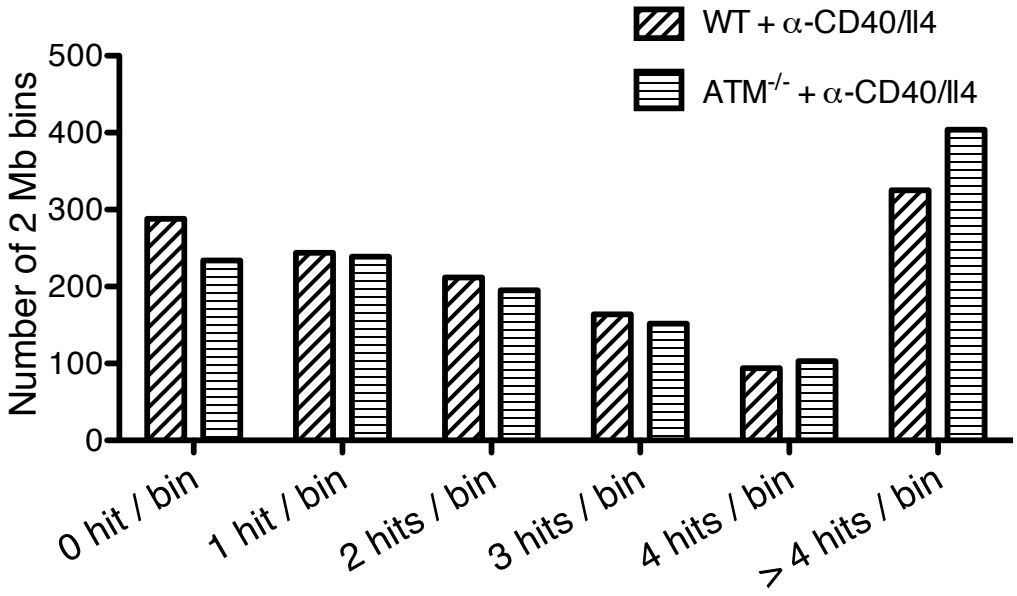


Figure 28. Statistical analysis of translocations distributed across the genome of in α-CD40/IL-4 stimulated WT and ATM-deficient B cells shown in Figure 26. The number of 2 Mb bins containing 1, 2, 3, or 4 junctions is represented on the X axis. The Y axis represents the numbers of 2Mb bins containing indicated number of junctions. The mouse genome was divided into 2 Mb bins, and the number of bins containing indicated hits was counted separately, excluding the 2Mb bin around the *c-myc* breaksite.

3.1.2. ATM Prevents The Formation Of Dicentrics

Given that ATM-deficient B cell lymphomas routinely harbor translocations involving *IgH*, we analyzed the pattern of HTGTS translocation junctions on chromosome 12. In WT cells, most of the junctions were found in the *IgH* locus, and specifically in the Sμ, Sγ1, and Sε regions (Figure **29**), as previously reported (Chiarle et al., 2011). In libraries from ATM-deficient α-CD40/IL4-stimulated B cells, more translocations mapped on the telomeric portion of chr12 (where the *IgH* locus lies) as compared to WT (Figure 26). However, the translocations not only involved *IgH* S regions, but also accumulated over a 30 Mb region downstream of *IgH*, towards the centromere of chromosome 12 (Figure 30). In the absence of ATM, RAG-initiated DSBs have been shown to persist throughout B cell development and can undergo extensive end resection (Callen et al., 2007). However, closer examination of the translocation junctions in our ATM-deficient translocation libraries revealed an even distribution of junctions in both (+) and (-) orientations (Figure 20), suggesting they originated from a different mechanism. In this regard, formation of chromosome 12 dicentrics following *IgH* locus breaks would lead to generation of secondary DSBs, with patterns compatible to the ones observed here (Figure 31).

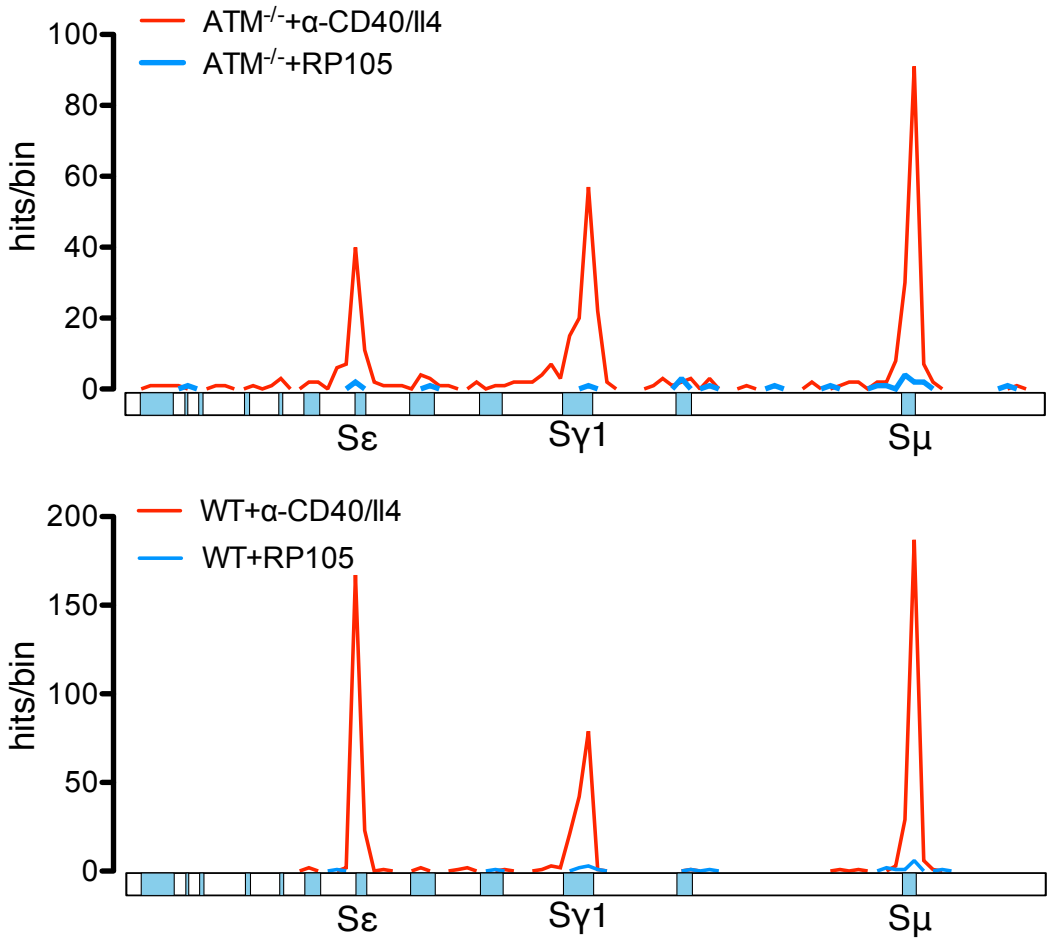


Figure 29. Sμ, Sγ1, and Sε regions are translocation hotspots in α-CD40/IL-4 stimulated but not RP105 stimulated B cells. Distributions of translocation junctions at the *IgH* locus in ATM$^{-/-}$ and WT stimulated B cells were shown respectively. Y-axis shows the number of hits in each bin, and the bin size is 2.5kb. A schematic showing gene structure of the *IgH* switch region is below each linear plot. The total number of each library was normalized to the cryptic I-SceI sites previously identified (Chiarle et al., 2012).

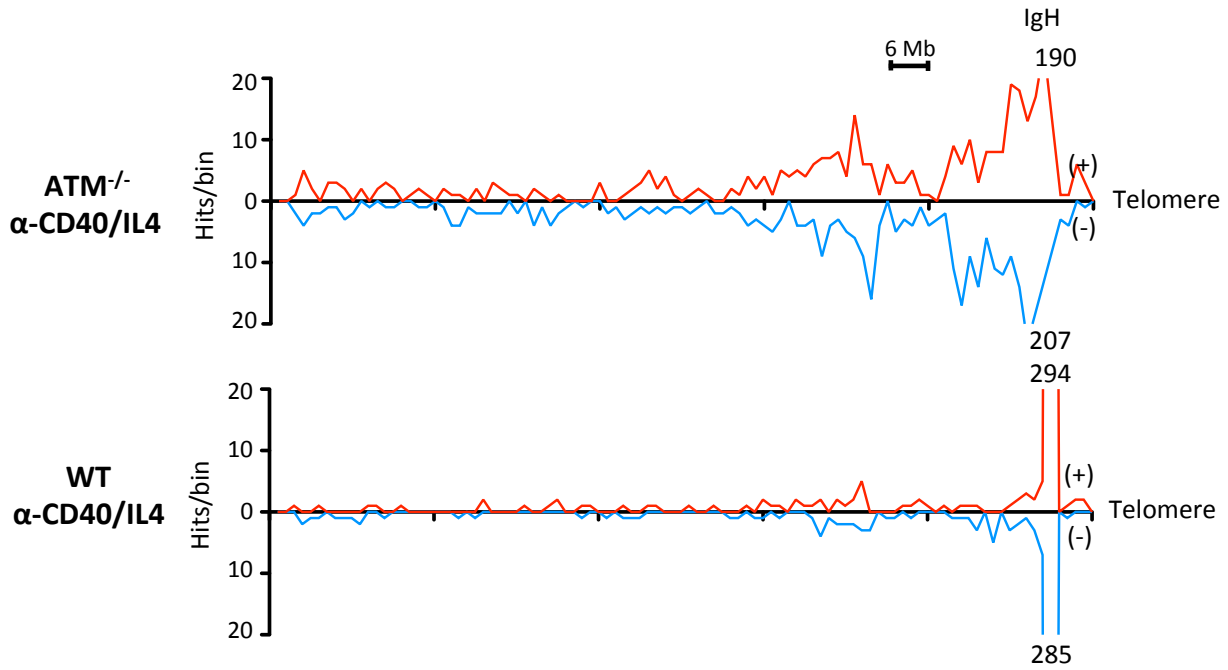


Figure 30. Distribution of translocation junctions on chromosome 12 from three independent HTGTS libraries of α-CD40/IL4-stimulated ATM$^{-/-}$ or WT B cells. The total translocation junction number from each library was normalized to number of junctions in cryptic I-SceI sites (see Material and Methods section for details). The bin size is 1.2Mb. The location of the *IgH* locus as well as the numbers of translocation junction within the S regions are indicated. The number of translocation junctions in (+) and (-) orientations are indicated.

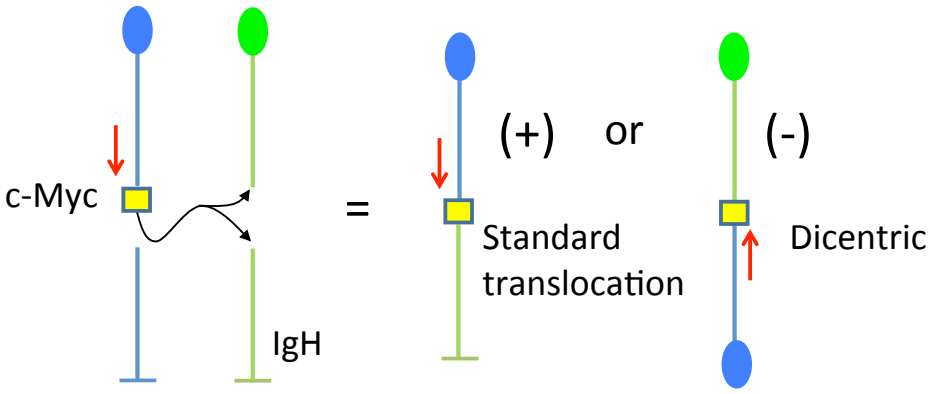


Figure 31. A model showing types of translocation junction that can be captured by out HTGTS assay. I-SceI-initiated DSB in *c-myc* can join to both ends of a new DSB downstream of the *IgH* locus. Translocation joining *c-myc* break to telomeric chromosome 12 is in (+) orientation while joining to centromeric chromosome 12 is in (-) orientation. The yellow box represents the *c-myc* gene. The red arrow indicates the position and direction of the HTGTS sequencing primer.

We next wanted to determine whether the formation of chromosome 12 dicentric was initiated solely by AID-dependent DSBs in culture (Franco et al., 2006) or also by RAG-initiated DSBs that persisted through B cell development. To this end, we generated HTGTS libraries from RP105-stimulated WT and ATM-deficient B cells. The RP105 stimulation allows B cells to survive and proliferate in culture without expressing AID or undergoing CSR, therefore eliminating AID-dependent DSBs in S regions. As expected, in RP105 stimulated cells, translocation junctions were barely found at S regions and other AID-dependent hot spots previously detected in α-CD40/IL4-stimulated B cells (Figure 29 and data not shown). However, the accumulation of junctions on the telomeric end of chromosome 12 was retained in libraries from RP105-stimulated ATM-deficient cells. Additionally, translocation junctions downstream of *IgH* were distributed in both (+) and (-) orientations as observed above. On the contrary, libraries from RP105-stimulated WT cells only showed a level of junctions in this region comparable to the one observed in α-CD40/IL4-stimulated WT B cells (Figure 32 and Figure 33). Taken together, these results indicate that RAG-initiated DSBs contribute to the formation of chromosome 12-12 dicentric which can provide a source of further DSBs in later stages of B cell development and that such DSBs can be substrates for chromosomal translocations.

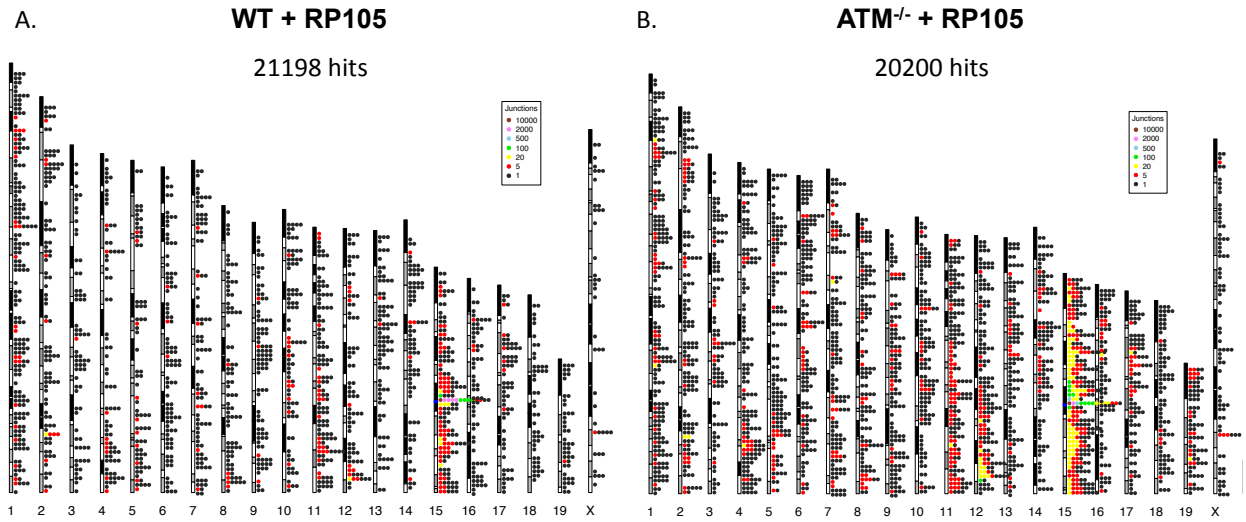


Figure 32. Distributions of genome-wide translocation in HTGTS libraries from RP105-stimulated (A) WT and (B) ATM-deficient B cells. Each plot represents a pool from three independent experiments. Genotype, treatment, and total number of junctions in each library are shown above each plot. The bin size is 2 Mb. Clusters of translocation hits are represented by color dots as showed in the square.

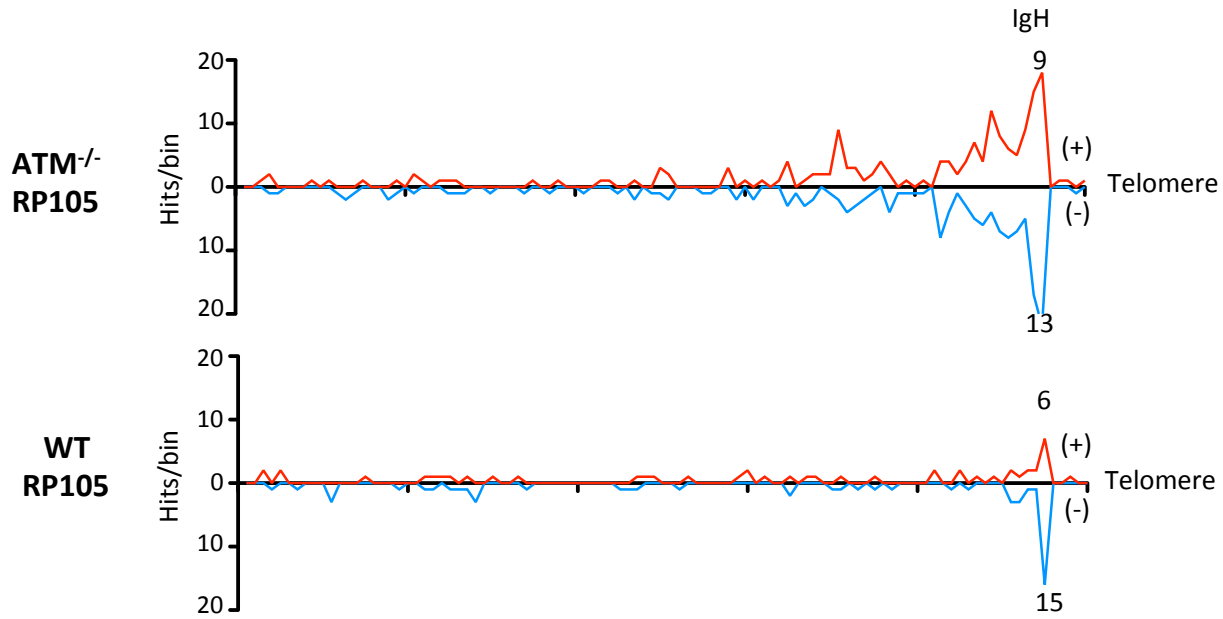


Figure 33. Distribution of translocation junctions on chromosome 12 from three independent HTGTS libraries of RP105-stimulated WT or ATM-deficient B cells. The total translocation junction number from each library was normalized to number of junctions in cryptic I-SceI sites (see Material and Methods section for details). The bin size is 1.2Mb. The location of the *IgH* locus as well as the numbers of translocation junction within the S regions are indicated. The number of translocation junctions in (+) and (-) orientations are indicated.

3.2. Discussion

3.2.1. RAG-Initiated DSBs May Persist Developmentally As Dicentrics And Give Rise To *de novo* DSBs In Mature B Cells.

To further elucidate mechanisms that could contribute to an ability of RAG-initiated and other types of DSBs to generate *c-myc* translocations in context of ATM-deficient mature B cells, we employed HTGTS to study the fate of DSBs introduced into *c-myc* in wild-type and ATM-deficient splenic B cells activated in culture under conditions that would or would not promote AID-activation. Mapping of genome-wide translocation junctions revealed that in both WT and ATM-deficient B cells *c-myc* DSBs frequently translocate in both the (+) and (-) chromosomal orientations to newly generated AID-initiated DSBs in donor and acceptor *IgH* S regions (Chiarle et al., 2011). However, in ATM-deficient B cells, but not wild-type B cells, *c-myc* translocation junctions also occurred at high levels throughout a 30 megabases region downstream (centromeric) of the *IgH* locus. This broad translocation hotspot region downstream of *IgH* in ATM-deficient cells contained a comparable distribution of translocation junctions in both (+) and (-) chromosomal orientations, indicating that they involved *de novo* generated DSBs downstream of *IgH* in activated ATM-deficient B cells. The possibility that these DSBs were generated by AID off-target activity over large region of chromosome 12 was conclusively eliminated by our finding that this region downstream of *IgH* was still a hotspot for DSBs and translocations to *c-myc* in ATM-deficient B cells cultured under conditions that do not lead to AID induction or *IgH* S region DSBs and translocations.

How could *de novo* DSBs be generated in such a large region downstream of *IgH* in ATM-deficient activated B cells? The origin of these aberrant forms of chromosome 12 has been

thought to involve unrepaired RAG-initiated *IgH* locus DSBs that persist through B cell development in the form of telomere-less chromosome 12 fragments and, by the time they are found in mature B cells, undergo resections that can delete large regions centromeric to the *IgH* locus (Callen et al., 2007). As a significant portion of the DSBs downstream of *IgH* in activated ATM-deficient B cells identified via HTGTS are AID-independent, they likely represent these proposed persistent RAG-initiated breaks, or at least a highly overlapping subset of them. However, a translocation mechanism based on persistence of broken telomere-less fragments of chromosome 12 could not explain our finding of equally high numbers of translocation junctions in both the (+) and (-) chromosomal orientations within the 30 Mb region downstream of *IgH*. We thus conclude that the generation of both (+) and (-) orientation junctions require the presence of DSB ends associated, respectively, with centromeric and telomeric ends of chromosome 12 (Figure 34).

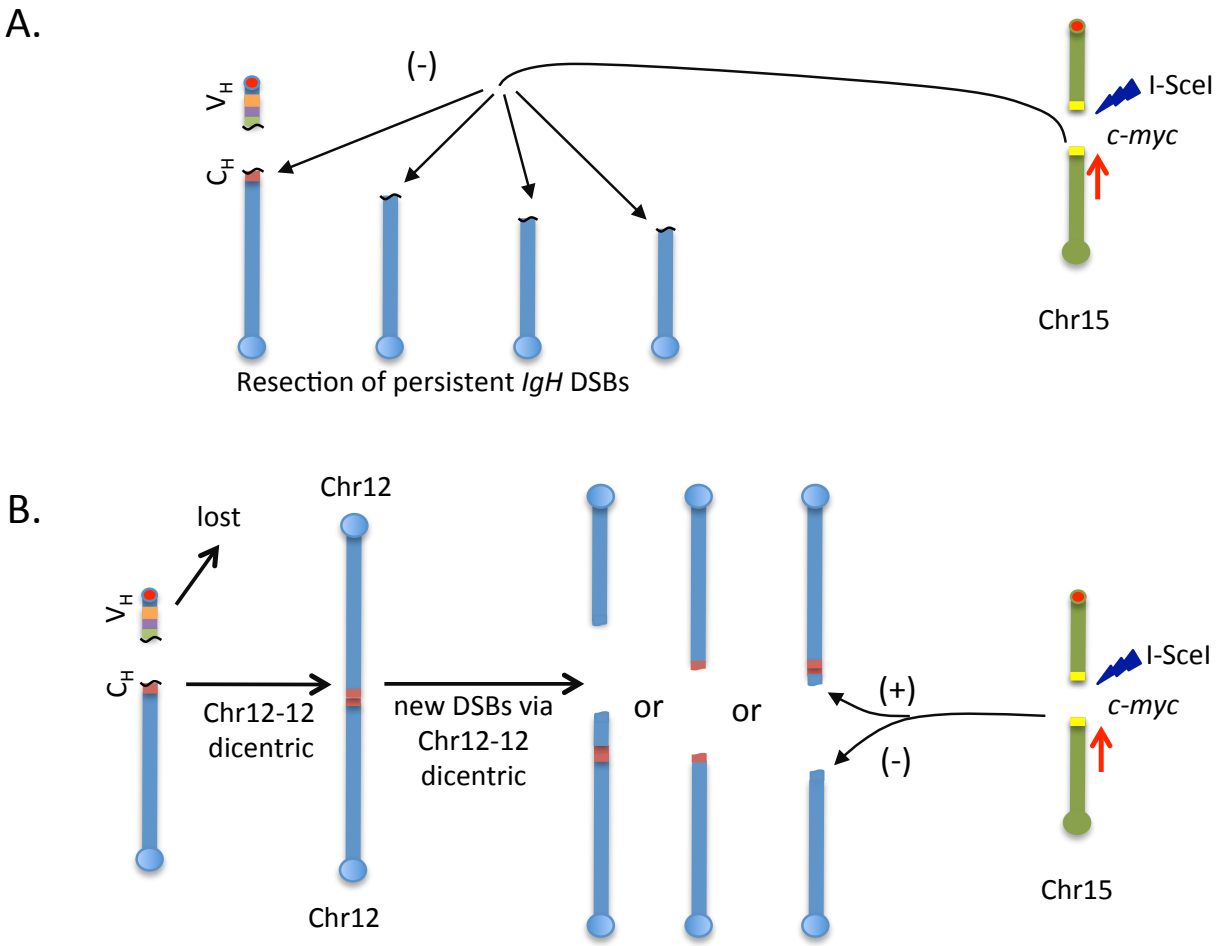


Figure 34. Two models for persistence of RAG-initiated DSBs in ATM-deficient B lymphocytes that provide translocation substrates in the peripheral B cells. (A) Persistence of *IgH* DSBs with resection due to combined defect in DNA repair and G1 checkpoint in the absence of ATM (Callen et al., 2007). (B) Persistence of *IgH* DSBs as dicentrics, resulting in generation of *de novo* DSBs in the periphery.

Franco and colleagues have shown that chromosome 12 dicentrics can be formed via aberrant repair of AID-initiated DSBs in activated ATM-deficient peripheral B cells (Franco et al., 2006). Our current findings now indicate that such dicentrics can and frequently do arise from AID-independent mechanisms, and can generate downstream DSBs that are substrates for translocations to *c-myc*. We propose that a large fraction of AID-independent DSBs within *IgH* and in the region far downstream of the locus in ATM-deficient peripheral B cells result from BFB cycles involving chromosome 12 dicentrics generated from unrepaired RAG-initiated *IgH* DSBs in pro-B cells by the mechanisms discussed above. Thus, our HTGTS studies reveal an unanticipated mechanism by which ATM suppresses the generation of chromosome breaks and translocations in peripheral B cells by preventing unrepaired RAG-initiated DSBs in pro-B cells from propagating as dicentrics during development which can further generate DSBs in peripheral B cells that can contribute to oncogenic translocations.

3.2.2. Implications For ATM-Deficient Mouse B And T Cell Lymphomas.

Together our studies indicate that, in certain genetic backgrounds, RAG-initiated *IgH* DSBs in ATM-deficient pro-B cells can result in the development of peripheral B cell lymphomas with complex *IgH* locus translocations, which lead to *c-myc* activation via gene amplification. Our findings support a model in which RAG-initiated dicentric translocations involving the *IgH* locus occur in pro-B cells and are propagated developmentally by a BFB mechanism. In this context, additional translocation and gene amplifications leading to the peripheral B cell lymphomas in the ATM-deficient background may occur at any stage of development including, as shown by our HTGTS studies, in mature B cells. Our findings also provide a model for the

generation of chromosomal aberrations found recurrently in ATM-deficient mouse T cell lymphomas. These tumors frequently harbor complex chromosome 14 translocations involving dicentrics generated during the *Tcrd* locus V(D)J recombination that lead to amplification of genes upstream of the *Tcrd* locus (Zha et al., 2010). Notably, these tumors also have characteristic T(12;14) translocations that fuse *Tcrd* breaks on chromosome 14 to chromosome 12 sequences that lie within a large region downstream of the *IgH* locus, which may result in a deletion of putative tumor suppressors (Zha et al., 2010). Based on our current findings, we propose that the substrates for translocations that lead to the large chromosome 12 translocations could be RAG-initiated DSBs in the JH locus, which are known to occur in developing T cells. In the absence of ATM, the RAG-initiated DSBs form dicentrics that continue to undergo BFB breaks and serve as substrates for translocation to persistent RAG-initiated *Tcrd* breaks.

Nearly all of the peripheral DA, AB, and DAB B cell lymphomas arise from peripheral IgM^+ B cells have not undergone either CSR or SHM. Given our finding that activated ATM-deficient B cells in culture form AID-initiated dicentrics and other translocations at a high frequency (Franco et al., 2006), one might expect that such AID-initiated DSBs would also give rise to oncogenic translocations. The lack of B cell lymphomas derived from ATM-deficient GC B cells might result from more strict checkpoints in GC or other AID-activated B lymphocytes, which eliminate cells carrying off-target AID DNA damage. In this context, perhaps the susceptible ATM-deficient IgM^+ B cell population is more tolerant of DNA damage and oncogene stress than *in vivo* ATM-deficient GC or CSR activated B cells. In addition, it is possible that progenitors for ATM-deficient IgM^+ B cell lymphomas arrive in the periphery with activated

c-myc genes providing proliferative advantages, allowing the cells to achieve full oncogenic transformation at this developmental stage.

Our findings could be relevant to certain pre-germinal center human B cell lymphomas. For example, a subset of mantle cell lymphomas (MCLs) derive from pre-germinal center B cells with ATM mutations (Stankovic et al., 2002). MCLs also frequently harbor translocations between chromosome 14 and chromosome 11 that lead to deregulation of *cyclinD1* gene through translocation to the *IgH* JH region. Moreover, a number of MCLs are documented as having complex chromosomal rearrangements with co-amplifed *IgH/cyclinD1* genes (Menanteau and Martinez-Climent, 2013; Setoodeh et al., 2013).

CHAPTER 4: FUTURE DIRECTIONS

The observation that the DJβ cassette, despite not being directly involved in translocations, renders the modified *c-myc* allele a more frequent translocation target is of substantial interest and warrants further analysis. In this regard, it is important to note that we have only studied one DJβ cassette targeted mouse line, which was used to generate mice in both the DA and DAB cohorts. We cannot therefore exclude the possibility that unknown genetic alterations in the targeted embryonic stem (ES) cell clone used to generate the c-myc$^{DJβ/WT}$ mice conferred predisposition for B cell lymphomas carrying *c-myc* translocations. To unequivocally rule out this possibility, several independently targeted c-myc$^{DJβ/WT}$ mouse lines would need to be generated and crossed into the ATM-deficient background to generate cohorts that would be monitored for B cell lymphoma development.

If the independently targeted DA and DAB mouse lines continued to develop mature B cell lymphomas with the characteristic *IgH/c-myc* amplification, it would indicate that the DJβ cassette plays a role in tumor development. If this was the case, it would be of great interest to determine if the oncogenic event is driven by the cassette is a RAG-independent or RAG-dependent process. In the latter context, we have hypothesized that the generation of DSBs downstream of *c-myc* might be linked to the presence of the DJβ cassette in *c-myc* intron 1. However, we cannot exclude contributions of an unknown mechanism, for example some form of oncogenic activation post amplification, by which the insertion of the DJβ cassette into *c-myc* intron that dictates cellular selection of oncogenic translocation/amplification of the *c-myc^{DJβ}* allele. We outline below experiments that would help distinguish these alternative possibilities.

To test the potential mechanisms by which the DJβ insertion contributes to oncogenic rearrangements/amplifications in the *c-myc* allele into which it is inserted, we could mutate the

RSSs within the DJβ cassette such that it cannot undergo V(D)J recombination. The RSS-mutated DJβ cassette would then be targeted into the same region in *c-myc* intron 1 as the previously targeted the c-myc^{DJβ} allele, generating the c-myc$^{mut\text{-}DJβ}$ allele. Mice containing the c-myc$^{mut\text{-}DJβ}$ allele would be crossed into ATM$^{-/-}$ and ATM$^{-/-}$-Eμ-Bcl2 background to generate mut-DA and mut-DAB mouse cohorts, respectively, that would be monitored for tumor development. If the mut-DA and mut-DAB mice developed B cell lymphomas with the same characteristics as DA and DAB mice, it would suggest that the oncogenic event is RAG-independent and more likely resulted from other mechanisms such contributing to further activation of the amplified c-myc alleles that were selected for subsequent translocations and initial amplifications. In this regard, the DJβ insertion itself is not sufficient to activate *c-myc* for tumorigenesis since all of the DA and DAB B cell lymphomas harbor translocations and amplifications.

In regards to cellular selection against translocations that disrupt *c-myc* intron 1, it should be noted that a translocation into the region where the cassette is targeted might lead to *c-myc* inactivation, via a disruption of an alternative promoter that lies upstream of *c-myc* exon 2 (Bernard et al., 1983; Saito et al., 1983). Such inactivation would obviate translocations directly into the first *c-myc* intron from contributing to *c-myc* oncogenic activation. The lack of *c-myc* activation via translocations into murine *c-myc* intron 1 is further demonstrated by studies of an Xrcc4/p53-deficient peripheral B cell lymphoma (CXP) mouse model. Thus, peripheral B cell lymphomas from this background harbor reciprocal T(12;15) and T(15;12) translocations with translocation junctions on T(15;12) involving the joining of sequences within or around *IgH* S region on chromosome 12 to sequences that consistently lie upstream of *c-myc* exon 1 (Wang et al., 2008). Based on these considerations, it is conceivable the lack of

translocations directly involving the DJβ cassette in our current experimental system would reflect cellular selection during tumorigenesis for B lymphocytes with breakpoints outside of *c-myc*, which would maintain the *c-myc* transcription start sites located within intron 1 (Stanton et al., 1983). Accordingly, the retention of transcription start site in *c-myc* intron 1 in human Burkitt's lymphoma T(8;14) translocation is shown to be essential in ectopic activation of *c-myc* by *IGH* regulatory elements in this cancer (Janz, 2006).

If the mut-DA and mut-DAB mouse cohorts did not succumb to the same type of mature B cell lymphomas development as in the DA and DAB mice, it would support the notion that the DJβ cassette contributes oncogenic event through an indirect RAG/V(D)J recombination associated mechanism. To test the hypothesis that aberrant V(D)J recombination into the DJβ cassette could, if in a different position than intron 1, contribute to B cell lymphoma in mice, we would generate ATM-deficient experimental mouse cohorts harboring the c-myc^{DJβ} allele with the DJβ cassette targeted to a region either upstream or downstream of *c-myc* instead of in intron one. Translocations upstream of *c-myc* should not inactivate the gene but could potentially lead to its ectopic activation by the *IgH* regulatory elements (Gostissa et al., 2009; Wang et al., 2008), while translocations downstream of *c-myc* could lead to dicentrics that would provide substrates for *c-myc* gene amplification via BFB mechanism as discussed.

If we do find evidence that the DJβ cassette promotes *c-myc* oncogenic translocations through an indirect V(D)J recombination associated mechanism, it would also be of great interest to test the hypothesis that the DJβ cassette promotes translocations downstream of the *c-myc* gene. In the discussion section, we proposes a potential mechanism that may involve increasing levels of DSBs in the downstream region, which then provide substrates for dicentric

translocations and oncogenic amplifications. In this context, we further speculate that the presence of the DJβ cassette in *c-myc* first intron increases the likelihood of RAG cleavage at cryptic RSSs downstream of *c-myc* by providing a paired functional RSS for 12/23 restricted cleavage of isolated cryptic RSSs downstream of *c-myc*. To test this hypothesis, we would design a PCR assay with one primer upstream of the Dβ segment and another primer downstream of cryptic RSSs that mapped near the translocation junctions identified in DA and DAB tumors. This strategy would allow us to detect the joining of RAG-initiated DJβ DSBs to the downstream cryptic RSSs. If the presence of the DJβ cassette increased the level of DSBs in the region downstream of *c-myc*, we would detect these joins in pro-B cells from pre-tumor DA mice.

Another characteristic of recurrent translocations in mature B cell lymphomas from the three experimental cohorts in this study is the joining of sequences downstream of *c-myc* to a small region in and around the IgH J_H segment cluster, thereby linking the *IgH* intronic enhancer (iEμ) to *c-myc*. Such juxtaposition may contribute to mature B cell lymphoma development in ATM-deficient background, especially in light of the observation that in some tumors (e.g. DAB473) low levels of *c-myc* amplification correlated with translocation breakpoints closer to the *c-myc* gene. To test the role of iEμ in ATM-deficient B cell tumor development, we would generate DAB mouse line that also lacks iEμ and observe the experimental mice for tumor development. If iEμ was dispensable for oncogenic transformation, the tumor types and characteristics would remain the same in mouse lines lacking iEμ. Of particular interest, if tumors still develop would be whether or not we observed increased *c-myc* amplification or

loss of frequent translocations close to *c-myc* since iEμ activity has been suggested to be limited over distance (Henderson and Calame, 1998).

It is of relevant that the mapping of translocation junctions by HTGTS in *in vitro* stimulated ATM-deficient splenic B cells isolated from c-myc$^{25xI\text{-}SceI}$ B cells identified a 30 Mb region downstream of the *IgH* locus on chromosome 12 to be a translocation hotspot. When stimulated under conditions that did not induce AID or CSR, this region was still a hotspot for DSBs and translocations to *c-myc.* We propose that a large fraction of these breakpoints results from BFB cycles involving chromosomes 12 dicentrics generated from unrepaired RAG-initiated *IgH* DSBs in pro-B cells. However, the finding from this experiment does not tell unequivocally if the source of the putative dicentrics that rise to the AID-independent downstream *IgH* DSBs in these experiments is RAG activity in pro-B cells. To unequivocally determine the potential roles of RAG in the formation of chromosome 12-12 dicentrics that lead to *de novo* DSBs downstream of *IgH* in mature B cells, we would generate HTGTS libraries from splenic B cells isolated from mice with dual ATM and RAG deficiency and B cells from mice with ATM, RAG, and AID deficiency that also harbor pre-assembled IgH and IgL chain variable region exons (HL) knock in alleles. These knock in alleles encode proteins that assemble to form a complete immunoglobulin molecule, thus allowing RAG-deficient B cells to progress through development (Lansford et al., 1998). If RAG-initiated DSBs were the major source of DSBs that form dicentrics which subsequently result in *de novo* DSBs in the 30 Mb region downstream of IgH in the absence of ATM, we would expect a very low level of translocation in this region in ATM and RAG plus AID deficient/HL compared to AID-deficient ATM-deficient/HL background.

CHAPTER 5: MATERIALS AND METHODS

Generation of DA mice. To introduce the c-myc$^{DJ\beta}$ allele (Ranganath et al., 2008) into the ATM-deficient background (Borghesani et al., 2000), we initially bred ATM$^{+/-}$ mice to mice homozygous for the c-myc$^{DJ\beta}$ allele. We then intercrossed the resulting ATM$^{+/-}$ c-myc$^{DJ\beta/wt}$ offspring, as well as second generation ATM$^{+/-}$ c-myc$^{DJ\beta/wt}$ and ATM$^{+/-}$ c-myc$^{DJ\beta/DJ\beta}$ mice, to obtain experimental ATM$^{-/-}$ c-myc$^{DJ\beta/DJ\beta}$ and ATM$^{-/-}$ c-myc$^{DJ\beta/wt}$ (DA) mice and control mice. Experimental animals were monitored for tumor development and killed for analysis when clear signs of disease appeared. All animal experiments were performed under protocols approved by the Institutional Animal Care and Use Committee of Boston Children's Hospital (Protocol 11-11-2074R).

Generation of AB mice. To introduce the Eμ-Bcl-2 transgene into ATM-deficient background, we bred ATM$^{+/-}$ mice to mice harboring Eμ-Bcl-2 transgene (Strasser et al., 1994). The resulting offspring (ATM$^{+/-}$Eμ-Bcl-2 mice) were then bred together to obtain experimental ATM$^{-/-}$ Eμ-Bcl-2 (AB) and control mice. Experimental animals were monitored for tumor development and killed for analysis when clear signs of disease appeared. All animal experiments were performed under protocols approved by the Institutional Animal Care and Use Committee of Boston Children's Hospital (Protocol 11-11-2074R).

Generation of DAB mice. To introduce both the c-myc$^{DJ\beta}$ allele and Eμ-Bcl-2 transgene into ATM-deficient background, we first generated c-myc$^{DJ\beta/wt}$ Eμ-Bcl-2 mice and then bred them to ATM$^{+/-}$ mice. The resulting ATM$^{+/-}$ c-myc$^{DJ\beta/wt}$ Eμ-Bcl-2 offspring were then bred to ATM$^{+/-}$ c-myc$^{DJ\beta/wt}$ to generate experimental cohort (DAB), comprising ATM$^{-/-}$ c-myc$^{DJ\beta/wt}$ Eμ-Bcl-2 mice

and $ATM^{-/-}$ $c\text{-}myc^{DJ\beta/DJ\beta}$ Eμ-Bcl-2 mice. $ATM^{-/-}$ $c\text{-}myc^{DJ\beta/wt}$, $ATM^{+/-}$ $c\text{-}myc^{DJ\beta/wt}$ Eμ-Bcl-2, and Eμ-Bcl-2 mice were kept as controls. Experimental animals were monitored for tumor development and killed for analysis when clear signs of disease appeared. All animal experiments were performed under protocols approved by the Institutional Animal Care and Use Committee of Boston Children's Hospital (Protocol 11-11-2074R).

Flow-cytometry Analysis. Single-cell suspensions from tumor masses and control organs were stained with CyChrome (CyC)-labeled anti-mouse B220 (eBiosciences), FITC-labeled anti-mouse CD43 (BD Biosciences), and RPE-labeled anti-mouse IgM (Southern Biotech) antibodies or with FITC-labeled anti-mouse Igκ (BD Biosciences) and PE-labeled anti-mouse Igλ (BD Biosciences). Data acquisition was performed on a FACSCalibur flow cytometer equipped with CellQuest software (Becton Dickinson). Analysis was performed with FlowJo software (Tree Star).

Immunohistochemical Analysis. Freshly isolated tumor tissues were fixed in 10% (vol/vol) buffered formalin overnight and stored in 70% (vol/vol) ethanol. Paraffin-embedded tissues were sectioned and stained individually with H&E, anti-mouse B220, anti-mouse CD3, anti-mouse pan-Ig, anti-mouse c-myc, and anti-mouse TdT. Histologic diagnoses were made on the basis of established criteria (Morse et al., 2002).

Southern Blotting. Genomic DNA was isolated from tumor masses or normal control tissues. Southern blotting was performed with 10 μg of genomic DNA digested overnight with appropriate restriction enzyme. Digested DNA was then size-separated on 0.8% agarose gel and

transferred to Zeta-Probe GT (Biorad) nylon membrane. Hybridization with ^{32}P-labeled DNA probe was performed in 50% (vol/vol) formamide/SScPE at 42°C overnight. The J_{H4-3} probe was a 1.6-kb HindIII/EcoRI fragment downstream of J_{H4}; the Jκ probe was a 1-kb fragment downstream of $J_{κ5}$; and the Cμ probe was a 869bp XbaI/BamHI fragment. The 3' c-myc probe was 2.5 kb XhoI/BamHI fragment.

Somatic Hypermutation Analysis. The genomic region encompassing J_{H1} to J_{H4} and part of the intron downstream of J_{H4} was PCR-amplified from tumor DNA using degenerate oligonucleotides corresponding to the different VH families as forward primers and oligonucleotides downstream of J_{H4} (5′ AGGCTCTGAGATCCCTAGACAG3′ or 5′ CCTCTCCAGTTTCGGCTGATCC3′) as reverse primers. Proofreading polymerase (iProof High-Fidelity DNA Polymerase, Biorad) was used for amplification and PCR conditions were as previously published (Ehlich et al., 1994). Amplification products were purified from agarose gel and submitted to sequencing. Sequences were compared with the published 129/Sv and C57/B6 sequences (accession nos. NT_114985.2 and NT_166318.1, respectively).

Cytogenetic Analysis. To generate metaphase, freshly isolated lymphomas were made into single-cell suspensions and cultured for 3-6 hours in RPMI medium containing 15% fetal calf serum and 50ng/mL Colcemid (KaryoMAX Colcemid Solution; GIBCO). Metaphase spreads were prepared according to standard protocols (Franco et al., 2006). Spectral karyotyping was performed with a mouse SKY paint kit (Applied Spectral Imaging) following manufacturer's indications. Images were acquired with BX61 Microscope (Olympus) equipped with a motorized

automatic stage, a cooled-CCD camera, and an interferometer (Applied Spectral Imaging). A 63× objective was used. Analysis was performed with the HiSKY software (Applied Spectral Imaging). At least 15 metaphases per each sample were analyzed. Two-color FISH assays for *IgH* and *c-myc* loci were performed as previously described with the indicated BAC probes (Franco et al., 2006). Metaphases were first hybridized with chromosome-specific fluorochrome-conjugated paints to mouse chromosome 12 and 15, following the manufacturer's instructions (Applied Spectral Imaging). After image acquisition as outlined for SKY experiments, selected metaphase slides were treated with 2xSSC solution for 20 minutes at room temperature to remove painting probes. Subsequently, the slides were dehydrated through serial ethanol washes and then re-hybridize with fluorochrome-labled BAC probes. To prepare BAC probes, BACs were labeled with either biotin (Biotin-Nick Translation Mix; Roche) or digoxigenin (Dig-Nick Translation Mix; Roche), as per manufacturer instructions. 200 ng of BAC DNA was precipitated with mouse Cot1 DNA (Invitrogen; ratio of BAC DNA to Cot1 DNA, 1:20), resuspended in 15 µl of hybridization solution (50% formamide, 2X SSC, 10% dextran sulfate, and 0.15% SDS), and co-denatured on slides for 4 min at 73°C. Slides were then incubated in humidified chamber at 37°C for 16 h, washed three times in 50% formamide/2X SSC for 5 min at 45°C, washed three times in 0.1X SSC for 5 min at 45°C, and incubated with avidin-Cy3 and anti–digoxigenin-FITC (Roche; 1:200 and 1:400 dilution, respectively) in 2XSSC/0.05% Tween-20 for 1 h at 37°C. After three washes in 4XSSC/0.05% Tween-20, slides were mounted in Vectashield with DAPI (Vector Laboratories). Image acquisition was performed as outlined above.

Cloning translocation junction. Genomic DNA from the tumor was individually digested with various restriction enzymes such as XbaI, SacI, and PstI for 12 hours and then purified with phenol–chloroform. Purified DNA was then end-repaired with a combination of T4 DNA polymerase, DNA pol I Klenow fragment, and T4 polynucleotide kinase for 30 minutes at 20°C. End-repaired DNA was purified and A-tailed with Klenow polymerase without 3' exonuclease activity. A-tailed fragments were then ligated to an asymmetric adaptor at a molar ratio of 20:1 for one hour at 25°C. Nested PCR with adaptor- and locus-specific PCR primers were performed on ligated and purified DNA fragment. The first round of PCR, 100-250 ng of DNA was amplified in a final volume of 50 µl using a forward primer downstream of *c-myc* locus and an adaptor-specific reverse primer (see table below for primer sequences) and Phusion polymerase (Finnzymes). 20 PCR cycles were performed in the following conditions: 98°C for 20 s, 65°C for 30 s, and 72°C for 4 minutes. The second round of PCR was performed with another set of primers nested from the first *c-myc* and adaptor-specific primers with the same condition as described above but with 30 PCR cycles (see table below for primer sequences). After amplification, PCR products were run on 1% agarose gel and the PCR band was cut out and DNA extracted from the gel. Purified PCR product was sent to sequencing. Sequencing reads from each tumor samples were aligned to mouse reference genome (Mouse July 2007-NCBI Build37/mm9) with the BLAT program (Kent, 2002).

Gene targeting of c-myc$^{25xI\text{-}SceI}$ allele. To generate the targeting construct, the I-SceI site cassette was inserted into the pLNTK targeting vector and flanked with a 4.6-kb SphI–SphI 5' homology arm encompassing *c-myc* exon 1, and a 2.6-kb SphI–SphI 3' homology arm containing

c-myc exons 2 and 3. The targeting construct was electroporated into TC1 (129/Sv) ES cells, and screened EcoRI-digested DNA to identify potentially targeted ES cell clones by Southern blotting with a 5' probe consisting of the 1.5-kb XbaI‑XbaI fragment upstream of *c-myc* exon 1, and a 3' probe consisting of the 1.5-kb XhoI‑KpnI fragment downstream of *c-myc* exon 3. The neomycin cassette was deleted from targeted ES clones by infection with adenovirus expressing Cre-recombinase and used those to inject for germline transmission to generate c-myc$^{25xI\text{-}SceI}$/WT mice. To generate the ROSA$^{I\text{-}SceI\text{-}GR}$ allele, the previously described strategy (Sasaki et al., 2006) was followed to target an I-SceI-GR/IRES-tdTomato expression cassette into TC1 (129/Sv) ES cells. In both cases, the Neo cassette was deleted from targeted ES clones by infection with a Cre-recombinase-expressing adenovirus before generating ROSA$^{I\text{-}SceI\text{-}GR}$ mice.

B cell culture conditions, retroviral infection, and CSR assays. CD43- B cells were isolated from spleen and were put in culture at a density of 1x10^6/ml in RPMI medium containing fetal bovine serum, αCD40 (1 μg/ml, eBioscience) and IL-4 (20 ng/ml, R&D Systems). Retroviral supernatants were prepared from Phoenix packaging cells transfected with pMX-I-SceI vector or empty pMX vector as control. After 24 hr of culture, retroviral infection was performed by adding one volume of viral supernatant, spinning for 1.5 hr at 2400 RPM at 24 °C in the presence of polybrene at 2.5 μg/ml and incubating cells with viral supernatant overnight. After infection, medium was changed and cells were replated at 0.5x10^6/ml. At day 4 of stimulation, infection efficiency was evaluated by measuring the percentage of cells expressing the retroviral IRES-GFP and was found to be usually between 50% and 85%. ROSA$^{I\text{-}SceI\text{-}GR}$ B cells were cultured in RPMI medium as above but with 15% charcoal-stripped fetal bovine serum to minimize

nonspecific activation of the I-SceI-GR fusion and thus obtain high levels of cutting at the appropriate time. CSR was evaluated in the case of retrovirally infected cells by staining with Cy5-PE-labeled anti-mouse B220 (eBiosciences) and PE-labeled anti-mouse IgG1 (BD Biosciences). The TA-induced ROSA$^{I\text{-}SceI\text{-}GR}$ B cells were evaluated by staining with Cy5-PE-labeled anti-mouse B220 and FITC-labeled anti-mouse IgG1 (BD Biosciences). CSR ranged between 15% and 40% for retrovirus-infected B cells and 9%–12% for TA-induced B cells, due to lower proliferation rates in charcoal stripped serum. DNA was prepared from cells at day 4 of culture by standard methods.

Generation of HTGTS libraries by adaptor-PCR. Genomic DNA from c-myc$^{25xI\text{-}SceI}$ and c-myc$^{25xI\text{-}SceI}$/ROSA$^{I\text{-}SceI\text{-}GR}$ B cells was digested overnight with HaeIII and HaeIII-generated blunt ends were A-tailed with Klenow polymerase. An asymmetric adaptor (composed of an upper linker and a lower 3'-modified linker) was then ligated to fragmented DNA at a molar ratio of 20:1 for 30 min at 25° C. To remove unrearranged I-SceI cassettes, ligation reactions were digested either with I-SceI, or with both EcoRV and XbaI. The blocking digests were carried out for 8 hr. Translocation junctions were then PCR-amplified using the emulsion (em)-PCR approach as previously described (Williams et al., 2006). In the first round of PCR, 1 µg of DNA was amplified in a final volume of 50 µl, using a biotinylated forward primer (Myc-H; see table below for primer sequences) and an adaptor-specific reverse primer (AP1) and Phusion polymerase (Finnzymes). 20 PCR cycles were performed in the following conditions: 98°C for 10 s, 58°C for 30 s, and 72°C for 30 s. Multiple reactions were performed in generating large-scale libraries. Thereafter, biotinylated PCR products were isolated using the Dynabeads MyOne Streptavidin

C1 kit (Invitrogen) and an additional 3 hr digestion with blocking enzymes was performed. PCR products were eluted from the beads via a 30 min incubation at 65°C in 95% formamide/10 mM EDTA and purified. The purified products were then amplified in a second round with em-PCR. 30 µl of the first round PCR product, 80 pmol each of primers AP2 and Myc-Prelox, 20 µl ultra-pure BSA (Ambion) and 2 µl Taq polymerase (QIAGEN) in a final volume of 200 µl were emulsified in 400 µl of an oil-surfactant mixture. The emulsion mixture was divided into 50 µl individual aliquots and PCR was performed using the following conditions: 30 cycles of 94°C for 30 s, 60°C for 30 s, and 72°C for 1 min. Following PCR, the products were pooled and centrifuged in a table-top centrifuge for 5 min at 14,000 RPM to separate the phases and the oil layer was removed. The sample was then extracted 3 times with 1 ml of water-saturated diethyl ether and DNA was re-purified. The third, non-emulsion, round of PCR was performed with the same primers as in round 2, but with the addition of linkers and barcodes for MiSeq sequencing. After amplification, the pooled PCR reactions were size-fractionated between 200 and 800 base pairs on agarose gel. PCR amplicons were sequenced by MiSeq (Illumina) platform. Raw sequences were analyzed and mapped as described (Chiarle et al., 2011).

BIBLIOGRAPHY

Ahnesorg, P., Smith, P. & Jackson, S. XLF interacts with the XRCC4-DNA ligase IV complex to promote DNA nonhomologous end-joining. *Cell* **124,** 301–313 (2006).

Alt, F., Zhang, Y., Meng, F., Guo, C. & Schwer, B. Mechanisms of programmed DNA lesions and genomic instability in the immune system. *Cell* **152,** 417–429 (2013).

Ambrose, M. & Gatti, R. Pathogenesis of ataxia-telangiectasia: the next generation of ATM functions. *Blood* **121,** 4036–4045 (2013).

Arnal, S., Holub, A., Salus, S. & Roth, D. Non-consensus heptamer sequences destabilize the RAG post-cleavage complex, making ends available to alternative DNA repair pathways. *Nucleic Acids Res.* **38,** 2944–2954 (2010).

Barlow, C. *et al.* Atm-deficient mice: a paradigm of ataxia telangiectasia. *Cell* **86,** 159–171 (1996).

Bassing, C. & Alt, F. The cellular response to general and programmed DNA double strand breaks. *DNA Repair* **3,** 781–796 (2004). [A]

Bassing, C. & Alt, F. H2AX may function as an anchor to hold broken chromosomal DNA ends in close proximity. *Cell Cycle* **3,** 149–153 (2004). [B]

Bassing, C. *et al.* Histone H2AX: a dosage-dependent suppressor of oncogenic translocations and tumors. *Cell* **114,** 359–370 (2003).

Bassing, C. *et al.* Increased ionizing radiation sensitivity and genomic instability in the absence of histone H2AX. *Proc. Natl. Acad. Sci. U.S.A.* **99,** 8173–8178 (2002).

Bassing, C. *et al.* Recombination signal sequences restrict chromosomal V(D)J recombination beyond the 12/23 rule. *Nature* **405,** 583–586 (2000).

Bednarski, J. & Sleckman, B. Integrated signaling in developing lymphocytes: the role of DNA damage responses. *Cell Cycle* **11,** 4129–4134 (2012).

Bender, C. *et al.* Cancer predisposition and hematopoietic failure in Rad50(S/S) mice. *Genes Dev.* **16,** 2237–2251 (2002).

Bernard, O., Cory, S., Gerondakis, S., Webb, E. & Adams, J. Sequence of the murine and human cellular myc oncogenes and two modes of myc transcription resulting from chromosome translocation in B lymphoid tumours. *The EMBO Journal* **2**, 2375-2383 (1983).

Boboila, C. *et al.* Alternative end-joining catalyzes class switch recombination in the absence of both Ku70 and DNA ligase 4. *J. Exp. Med.* **207,** 417–427 (2010).

Boboila, C., Alt, F. & Schwer, B. Classical and alternative end-joining pathways for repair of lymphocyte-specific and general DNA double-strand breaks. *Adv. Immunol.* **116,** 1–49 (2012).

Bohgaki, T., Bohgaki, M. & Hakem, R. DNA double-strand break signaling and human disorders. *Genome Integr.* **1**, 15 (2010).

Bolland, D. *et al.* Antisense intergenic transcription in V(D)J recombination. *Nat. Immunol.* **5,** 630–637 (2004).

Borghesani, P. *et al.* Abnormal development of Purkinje cells and lymphocytes in Atm mutant mice. *Proc. Natl. Acad. Sci. U.S.A.* **97,** 3336–3341 (2000).

Bossen, C., Mansson, R. & Murre, C. Chromatin topology and the regulation of antigen receptor assembly. *Annu. Rev. Immunol.* **30,** 337–356 (2012).

Boultwood, J. Ataxia telangiectasia gene mutations in leukaemia and lymphoma. *J. Clin. Pathol.* **54,** 512–516 (2001).

Brady, B., Steinel, N. & Bassing, C. Antigen receptor allelic exclusion: an update and reappraisal. *J. Immunol.* **185,** 3801–3808 (2010).

Bredemeyer, A. *et al.* ATM stabilizes DNA double-strand-break complexes during V(D)J recombination. *Nature* **442,** 466–470 (2006).

Buck, D. *et al.* Cernunnos, a novel nonhomologous end-joining factor, is mutated in human immunodeficiency with microcephaly. *Cell* **124,** 287–299 (2006).

Bunting, S. & Nussenzweig, A. End-joining, translocations and cancer. *Nature Rev. Cancer* **13**, 443-454 (2013).

Butterworth, S. & Taylor, A. A subpopulation of t(2;14)(p11;q32) cells in ataxia telangiectasia B lymphocytes. *Hum. Genet.* **73,** 346–349 (1986).

Callebaut, I. *et al.* Cernunnos interacts with the XRCC4 x DNA-ligase IV complex and is homologous to the yeast nonhomologous end-joining factor Nej1. *J. Biol. Chem.* **281,** 13857–13860 (2006).

Callén, E. *et al.* Chimeric IgH-TCRalpha/delta translocations in T lymphocytes mediated by RAG. *Cell Cycle* **8,** 2408–2412 (2009).

Callén, E., Nussenzweig, M. & Nussenzweig, A. Breaking down cell cycle checkpoints and DNA repair during antigen receptor gene assembly. *Oncogene* **26,** 7759–7764 (2007).

Casellas, R., Yamane, A., Kovalchuk, A. & Potter, M. Restricting activation-induced cytidine deaminase tumorigenic activity in B lymphocytes. *Immunology* **126,** 316–328 (2009).

Cavalli, G. & Misteli, T. Functional implications of genome topology. *Nat. Struct. Mol. Biol.* **20,** 290–299 (2013).

Celeste, A. *et al.* Genomic instability in mice lacking histone H2AX. *Science* **296,** 922–927 (2002).

Celeste, A. *et al.* H2AX haploinsufficiency modifies genomic stability and tumor susceptibility. *Cell* **114,** 371–383 (2003).

Chaudhuri, J. *et al.* Evolution of the immunoglobulin heavy chain class switch recombination mechanism. *Adv. Immunol.* **94,** 157–214 (2007).

Chaudhuri, J., Khuong, C. & Alt, F. Replication protein A interacts with AID to promote deamination of somatic hypermutation targets. *Nature* **430,** 992–998 (2004).

Chen, J., Cooper, D., Férec, C., Kehrer-Sawatzki, H. & Patrinos, G. Genomic rearrangements in inherited disease and cancer. *Semin. Cancer Biol.* **20,** 222–233 (2010).

Chiarle, R. *et al.* Genome-wide Translocation Sequencing Reveals Mechanisms of Chromosome Breaks and Rearrangements in B Cells. *Cell* **147,** 107–119 (2011).

Cobb, R., Oestreich, K., Osipovich, O. & Oltz, E. Accessibility control of V(D)J recombination. *Adv. Immunol.* **91,** 45–109 (2006).

Corneo, B. *et al.* Rag mutations reveal robust alternative end joining. *Nature* **449,** 483–486 (2007).

Coster, G. & Goldberg, M. The cellular response to DNA damage: a focus on MDC1 and its interacting proteins. *Nucleus* **1,** 166–178 (2010).

Cowell, L., Davila, M., Kepler, T. & Kelsoe, G. Identification and utilization of arbitrary correlations in models of recombination signal sequences. *Genome Biol.* **3**, 72.1-72.20 (2002).

Crouch, E. *et al.* Regulation of AID expression in the immune response. *J. Exp. Med.* **204,** 1145–1156 (2007).

Davey, M. *et al.* Juxtaposition of the T-cell receptor alpha-chain locus (14q11) and a region (14q32) of potential importance in leukemogenesis by a 14;14 translocation in a patient with T-

cell chronic lymphocytic leukemia and ataxia-telangiectasia. *Proc. Natl. Acad. Sci. U.S.A.* **85,** 9287-9291 (1988).

de Yébenes, V. *et al.* miR-181b negatively regulates activation-induced cytidine deaminase in B cells. *J. Exp. Med.* **205,** 2199–2206 (2008).

Derheimer, F. & Kastan, M. Multiple roles of ATM in monitoring and maintaining DNA integrity. *FEBS Lett.* **584,** 3675–3681 (2010).

Di Noia, J. & Neuberger, M. Molecular mechanisms of antibody somatic hypermutation. *Annu. Rev. Biochem.* **76,** 1–22 (2007).

Difilippantonio, M. *et al.* Evidence for replicative repair of DNA double-strand breaks leading to oncogenic translocation and gene amplification. *J. Exp. Med.* **196,** 469-480 (2002).

Dik, W. *et al.* Different chromosomal breakpoints impact the level of LMO2 expression in T-ALL. *Blood* **110,** 388–392 (2007).

Dixon, J. *et al.* Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485,** 376–380 (2012).

Donehower, L. *et al.* Mice deficient for p53 are developmentally normal but susceptible to spontaneous tumours. *Nature* **356,** 215–221 (1992).

Dorsett, Y. *et al.* MicroRNA-155 suppresses activation-induced cytidine deaminase-mediated Myc-Igh translocation. *Immunity* **28,** 630–638 (2008).

Dudley, D., Chaudhuri, J., Bassing, C. & Alt, F. Mechanism and control of V(D)J recombination versus class switch recombination: similarities and differences. *Adv. Immunol.* **86,** 43–112 (2005).

Dvorak, C. & Cowan, M. Radiosensitive severe combined immunodeficiency disease. *Immunol. Allergy Clin. North Am.* **30,** 125–142 (2010).

Ebert, A. *et al.* The distal V(H) gene cluster of the Igh locus contains distinct regulatory elements with Pax5 transcription factor-dependent activity in pro-B cells. *Immunity* **34,** 175–187 (2011).

Ehlich, A., Martin, V., Müller, W. & Rajewsky, K. Analysis of the B-cell progenitor compartment at the level of single cells. *Curr. Biol.* **4,** 573–583 (1994).

Elson, A. *et al.* Pleiotropic defects in ataxia-telangiectasia protein-deficient mice. *Proc. Natl. Acad. Sci. U.S.A.* **93,** 13084–13089 (1996).

Evans, P. *et al.* Radiation-induced delayed cell death in a hypomorphic Artemis cell line. *Hum. Mol. Genet.* **15,** 1303–1311 (2006).

Feldhahn, N. *et al.* Activation-induced cytidine deaminase acts as a mutator in BCR-ABL1-transformed acute lymphoblastic leukemia cells. *J. Exp. Med.* **204,** 1157–1166 (2007).

Fernandez-Capetillo, O., Lee, A., Nussenzweig, M. & Nussenzweig, A. H2AX: the histone guardian of the genome. *DNA Repair* **3,** 959–967 (2004).

Franco, S., Alt, F. & Manis, J. Pathways that suppress programmed DNA breaks from progressing to chromosomal breaks and translocations. *DNA Repair* **5,** 1030–1041 (2006).

Frank, K. *et al.* Late embryonic lethality and impaired V(D)J recombination in mice lacking DNA ligase IV. *Nature* **396,** 173–177 (1998).

Gao, Y. *et al.* A targeted DNA-PKcs-null mutation reveals DNA-PK-independent functions for KU in V(D)J recombination. *Immunity* **9,** 367–376 (1998).

Gao, Y. *et al.* Interplay of p53 and DNA-repair protein XRCC4 in tumorigenesis, genomic stability and development. *Nature* **404,** 897–900 (2000).

Geisberger, R., Rada, C. & Neuberger, M. The stability of AID and its function in class-switching are critically sensitive to the identity of its nuclear-export sequence. *Proc. Natl. Acad. Sci. U.S.A.* **106,** 6736–6741 (2009).

Gennery, A. Primary immunodeficiency syndromes associated with defective DNA double-strand break repair. *Br. Med. Bull.* **77-78,** 71–85 (2006).

Giunta, S. & Jackson, S. Give me a break, but not in mitosis: the mitotic DNA damage response marks DNA double-strand breaks with early signaling events. *Cell Cycle* **10,** 1215–1221 (2011).

Goedecke, W., Eijpe, M., Offenberg, H., van Aalderen, M. & Heyting, C. Mre11 and Ku70 interact in somatic cells, but are differentially expressed in early meiosis. *Nat. Genet.* **23,** 194–198 (1999).

Gostissa, M. *et al.* Conditional inactivation of p53 in mature B cells promotes generation of nongerminal center-derived B-cell lymphomas. *Proc. Natl. Acad. Sci. U.S.A.* **110,** 2934–2939 (2013).

Gostissa, M. *et al.* Long-range oncogenic activation of Igh-c-myc translocations by the Igh 3' regulatory region. *Nature* **462,** 803–807 (2009).

Gostissa, M., Alt, F. & Chiarle, R. Mechanisms that promote and suppress chromosomal translocations in lymphocytes. *Annu. Rev. Immunol.* **29,** 319–350 (2011).

Gottlieb, T. & Jackson, S. The DNA-dependent protein kinase: requirement for DNA ends and association with Ku antigen. *Cell* **72,** 131–142 (1993).

Grabarz, A., Barascu, A., Guirouilh-Barbat, J. & Lopez, B. Initiation of DNA double strand break repair: signaling and single-stranded resection dictate the choice between homologous recombination, non-homologous end-joining and alternative end-joining. *Am. J. Cancer Res.* **2,** 249–268 (2012).

Gu, Y., Jin, S., Gao, Y., Weaver, D. & Alt, F. Ku70-deficient embryonic stem cells have increased ionizing radiosensitivity, defective DNA end-binding activity, and inability to support V(D)J recombination. *Proc. Natl. Acad. Sci. U.S.A.* **94,** 8076–8081 (1997).

Guipaud, O. *et al.* B-cell chronic lymphocytic leukaemia: a polymorphic family unified by genomic features. *Lancet Oncol.* **4,** 505–514 (2003).

Guirouilh-Barbat, J. *et al.* Impact of the KU80 pathway on NHEJ-induced genome rearrangements in mammalian cells. *Mol. Cell* **14,** 611–623 (2004).

Guirouilh-Barbat, J., Huck, S. & Lopez, B. S-phase progression stimulates both the mutagenic KU-independent pathway and mutagenic processing of KU-dependent intermediates, for nonhomologous end joining. *Oncogene* **27,** 1726–1736 (2008).

Gumy-Pause, F., Wacker, P. & Sappino, A. ATM gene and lymphoid malignancies. *Leukemia* **18,** 238–242 (2004).

Guo, C. *et al.* CTCF-binding elements mediate control of V(D)J recombination. *Nature* **477,** 424–430 (2011).

Gutierrez, A. *et al.* The BCL11B tumor suppressor is mutated across the major molecular subtypes of T-cell acute lymphoblastic leukemia. *Blood* **118,** 4169–4173 (2011).

Hecht, J. & Aster, J. Molecular biology of Burkitt's lymphoma. *J. Clin. Oncol.* **18,** 3707–3721 (2000).

Heintzman, N. *et al.* Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nat. Genet.* **39,** 311–318 (2007).

Helmink, B. & Sleckman, B. The response to and repair of RAG-mediated DNA double-strand breaks. *Annu. Rev. Immunol.* **30,** 175–202 (2012).

Henderson, A. & Calame, K. Transcriptional regulation during B cell development. *Annu. Rev. Immunol.* **16**, 163-200 (1998).

Hesslein, D. & Schatz, D. Factors and forces controlling V(D)J recombination. *Adv. Immunol.* **78,** 169–232 (2001).

Hewitt, S., Chaumeil, J. & Skok, J. Chromosome dynamics and the regulation of V(D)J recombination. *Immunol. Rev.* **237,** 43–54 (2010).

Hoeijmakers, J. DNA damage, aging, and cancer. *N. Engl. J. Med.* **361,** 1475–1485 (2009).

Huang, C. *et al.* Defects in coding joint formation in vivo in developing ATM-deficient B and T lymphocytes. *J. Exp. Med.* **204,** 1371–1381 (2007).

Jabbour, E. & Kantarjian, H. Chronic myeloid leukemia: 2012 update on diagnosis, monitoring, and management. *Am. J. Hematol.* **87,** 1037–1045 (2012).

Jacks, T. *et al.* Tumor spectrum analysis in p53-mutant mice. *Curr. Biol.* **4,** 1–7 (1994).

Jackson, K., Kidd, M., Wang, Y. & Collins, A. The shape of the lymphocyte receptor repertoire: lessons from the B cell receptor. *Front. Immunol.* **4,** 263 (2013).

Janz, S. Myc translocations in B cell and plasma cell neoplasms. *DNA Repair* **5,** 1213–1224 (2006).

Ji, Y. *et al.* The in vivo pattern of binding of RAG1 and RAG2 to antigen receptor loci. *Cell* **141,** 419–431 (2010).

Jones, J. & Simkus, C. The roles of the RAG1 and RAG2 “non-core” regions in V(D)J recombination and lymphocyte development. *Arch. Immunol. Ther. Exp. (Warsz).* **57,** 105–116 (2009).

Jung, D. *et al.* Extrachromosomal recombination substrates recapitulate beyond 12/23 restricted VDJ recombination in nonlymphoid cells. *Immunity* **18,** 65–74 (2003).

Jung, D., Giallourakis, C., Mostoslavsky, R. & Alt, F. Mechanism and control of V(D)J recombination at the immunoglobulin heavy chain locus. *Annu. Rev. Immunol.* **24,** 541–570 (2006).

Keim, C., Kazadi, D., Rothschild, G. & Basu, U. Regulation of AID, the B-cell genome mutator. *Genes Dev.* **27,** 1–17 (2013).

Kent, W. BLAT--the BLAST-like alignment tool. *Genome Res.* **12,** 656–664 (2002).

Kirsch, I., Brown, J., Lawrence, J., Korsmeyer, S. & Morton, C. Translocations that highlight chromosomal regions of differentiated activity. *Cancer Genet. Cytogenet.* **18,** 159–171 (1985).

Klein, I. *et al.* Translocation-capture sequencing reveals the extent and nature of chromosomal rearrangements in B lymphocytes. *Cell* **147,** 95–106 (2011).

Klein, U. & Dalla-Favera, R. New insights into the pathogenesis of chronic lymphocytic leukemia. *Semin. Cancer Biol.* **20,** 377–383 (2010).

Krangel, M. Mechanics of T cell receptor gene rearrangement. *Curr. Opin. Immunol.* **21,** 133–139 (2009).

Küppers, R. & Dalla-Favera, R. Mechanisms of chromosomal translocations in B cell lymphomas. *Oncogene* **20,** 5580–5594 (2001).

Küppers, R. Mechanisms of B-cell lymphoma pathogenesis. *Nat. Rev. Cancer* **5,** 251–262 (2005).

Lansford, R., Manis, J., Sonoda, E., Rajewsky, K. & Alt, F. Ig heavy chain class switching in Rag-deficient mice. *Int. Immunol.* **10,** 325–332 (1998).

Larmonie, N. *et al.* Breakpoint sites disclose the role of the V(D)J recombination machinery in the formation of T-cell receptor (TCR) and non-TCR associated aberrations in T-cell acute lymphoblastic leukemia. *Haematologica* **98,** 1173–1184 (2013).

Le Noir, S. *et al.* Extensive molecular mapping of TCRα/δ- and TCRβ-involved chromosomal translocations reveals distinct mechanisms of oncogene activation in T-ALL. *Blood* **120,** 3298–3309 (2012).

Lewis, S., Agard, E., Suh, S. & Czyzyk, L. Cryptic signals and the fidelity of V(D)J joining. *Mol. Cell. Biol.* **17,** 3125–3136 (1997).

Li, G. *et al.* Lymphocyte-specific compensation for XLF/cernunnos end-joining functions in V(D)J recombination. *Mol. Cell* **31,** 631–640 (2008).

Li, Z. *et al.* The XRCC4 gene encodes a novel protein involved in DNA double-strand break repair and V(D)J recombination. *Cell* **83,** 1079–1089 (1995).

Liao, M. *et al*. No requirement for V(D)J recombination in p53-deficient thymic lymphoma. *Mol Cell Biol.* **18**, 3495-3501 (1998).

Lieber, M. The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. *Annu. Rev. Biochem.* **79,** 181–211 (2010).

Lieberman-Aiden, E. *et al.* Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* **326,** 289–293 (2009).

Liu, M. *et al.* Two levels of protection for the B cell genome during somatic hypermutation. *Nature* **451,** 841–845 (2008).

Löbrich, M. & Jeggo, P. Harmonising the response to DSBs: a new string in the ATM bow. *DNA Repair* **4,** 749–759 (2005).

Lou, Z. *et al.* MDC1 maintains genomic stability by participating in the amplification of ATM-dependent DNA damage signals. *Mol. Cell* **21,** 187–200 (2006).

Luning Prak, E., Monestier, M. & Eisenberg, R. B cell receptor editing in tolerance and autoimmunity. *Ann. N. Y. Acad. Sci.* **1217,** 96–121 (2011).

Ma, Y., Pannicke, U., Schwarz, K. & Lieber, M. Hairpin opening and overhang processing by an Artemis/DNA-dependent protein kinase complex in nonhomologous end joining and V(D)J recombination. *Cell* **108,** 781–794 (2002).

Magrath, I. Epidemiology: clues to the pathogenesis of Burkitt lymphoma. *Br. J. Haematol.* **156,** 744–756 (2012).

Mahowald, G. *et al.* Aberrantly resolved RAG-mediated DNA breaks in Atm-deficient lymphocytes target chromosomal breakpoints in cis. *Proc. Natl. Acad. Sci. U.S.A.* **106,** 18339–18344 (2009).

Mailand, N. *et al.* RNF8 ubiquitylates histones at DNA double-strand breaks and promotes assembly of repair proteins. *Cell* **131,** 887–900 (2007).

Manis, J. *et al.* 53BP1 links DNA damage-response pathways to immunoglobulin heavy chain class-switch recombination. *Nat. Immunol.* **5,** 481–487 (2004).

Mao, C. *et al.* T cell-independent somatic hypermutation in murine B cells with an immature phenotype. *Immunity* **20,** 133–144 (2004).

Matei, I. *et al.* ATM deficiency disrupts Tcra locus integrity and the maturation of CD4+CD8+ thymocytes. *Blood* **109,** 1887–1896 (2007).

Matei, I., Guidos, C. & Danska, J. ATM-dependent DNA damage surveillance in T-cell development and leukemogenesis: the DSB connection. *Immunol. Rev.* **209,** 142–158 (2006).

Matheson, L. & Corcoran, A. Local and global epigenetic regulation of V(D)J recombination. *Curr. Top. Microbiol. Immunol.* **356,** 65–89 (2012).

Matthews, A. & Oettinger, M. RAG: a recombinase diversified. *Nat. Immunol.* **10,** 817–821 (2009).

Matthews, A. *et al.* RAG2 PHD finger couples histone H3 lysine 4 trimethylation with V(D)J recombination. *Nature* **450,** 1106–1110 (2007).

McDonnell, T. *et al.* bcl-2-immunoglobulin transgenic mice demonstrate extended B cell survival and follicular lymphoproliferation. *Cell* **57,** 79–88 (1989).

McKinnon, P. ATM and the molecular pathogenesis of ataxia telangiectasia. *Annu. Rev. Pathol.* **7,** 303–321 (2012).

Meaburn, K., Misteli, T. & Soutoglou, E. Spatial genome organization in the formation of chromosomal translocations. *Semin. Cancer Biol.* **17,** 80–90 (2007).

Meek, D. Tumour suppression by p53: a role for the DNA damage response? *Nat. Rev. Cancer* **9,** 714–723 (2009).

Menanteau, M. & Martinez-Climent, J. Genomic profiling of mantle cell lymphoma. *Methods Mol. Biol.* **973,** 147–163 (2013).

Metcalfe, J., Heppell-Parton, A., McConville, C. & Taylor, A. Characterization of a B-lymphocyte t(2;14) (p11;q32) translocation from an ataxia telangiectasia patient conferring a proliferative advantage on cells in vitro. *Cytogenet. Cell Genet.* **56,** 91–98 (1991).

Minegishi, M. *et al.* Functional and molecular characteristics of acute lymphoblastic leukemia cells with a mature T-cell phenotype from a patient with ataxia telangiectasia. *Leukemia* **5,** 88–89 (1991).

Mitelman, F., Johansson, B. & Mertens, F. The impact of translocations and gene fusions on cancer causation. *Nat. Rev. Cancer* **7,** 233–245 (2007).

Morales, J. *et al.* 53BP1 and p53 synergize to suppress genomic instability and lymphomagenesis. *Proc. Natl. Acad. Sci. U.S.A.* **103,** 3310–3315 (2006).

Morse, H. *et al.* Bethesda proposals for classification of lymphoid neoplasms in mice. *Blood* **100,** 246–258 (2002).

Moshous, D. *et al.* Artemis, a novel DNA double-strand break repair/V(D)J recombination protein, is mutated in human severe combined immune deficiency. *Cell* **105,** 177–186 (2001).

Moshous, D. *et al.* Partial T and B lymphocyte immunodeficiency and predisposition to lymphoma in patients with hypomorphic mutations in Artemis. *J. Clin. Invest.* **111,** 381–387 (2003).

Muramatsu, M. *et al.* Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. *Cell* **102,** 553–563 (2000).

Neves, H., Ramos, C., da Silva, M., Parreira, A. & Parreira, L. The nuclear topography of ABL, BCR, PML, and RARalpha genes: evidence for gene proximity in specific phases of the cell cycle and stages of hematopoietic differentiation. *Blood* **93,** 1197–1207 (1999).

Nishana, M. & Raghavan, S. Role of recombination activating genes in the generation of antigen receptor diversity and beyond. *Immunology* **137,** 271–281 (2012).

Nussenzweig, A. & Nussenzweig, M. Origin of chromosomal translocations in lymphoid cancer. *Cell* **141,** 27–38 (2010).

O'Driscoll, M. & Jeggo, P. The role of double-strand break repair - insights from human genetics. *Nat. Rev. Genet.* **7,** 45–54 (2006).

Parada, L., McQueen, P. & Misteli, T. Tissue-specific spatial organization of genomes. *Genome Biol.* **5,** R44 (2004).

Parada, L., McQueen, P., Munson, P. & Misteli, T. Conservation of relative chromosome positioning in normal and cancer cells. *Curr. Biol.* **12,** 1692–1697 (2002).

Pasqualucci, L. *et al.* Hypermutation of multiple proto-oncogenes in B-cell diffuse large-cell lymphomas. *Nature* **412,** 341–346 (2001).

Pasqualucci, L. *et al.* Molecular pathogenesis of non-Hodgkin's lymphoma: the role of Bcl-6. *Leuk. Lymphoma* **44 Suppl 3,** S5–12 (2003).

Perlman, S., Boder, E., Sedgewick, R. & Gatti, R. A. Ataxia-telangiectasia. *Handb. Clin. Neurol.* **103,** 307–332 (2012).

Perlot, T. & Alt, F. Cis-regulatory elements and epigenetic changes control genomic rearrangements of the IgH locus. *Adv. Immunol.* **99,** 1–32 (2008).

Phillips, J. & Corces, V. CTCF: master weaver of the genome. *Cell* **137,** 1194–1211 (2009).

Pileri, S. & Falini, B. Mantle cell lymphoma. *Haematologica* **94,** 1488–1492 (2009).

Puebla-Osorio, N. & Zhu, C. DNA damage and repair during lymphoid development: antigen receptor diversity, genomic integrity and lymphomagenesis. *Immunol. Res.* **41,** 103–122 (2008).

Raghavan, S., Swanson, P., Wu, X., Hsieh, C. & Lieber, M. A non-B-DNA structure at the Bcl-2 major breakpoint region is cleaved by the RAG complex. *Nature* **428,** 88–93 (2004).

Ramiro, A. *et al.* Role of genomic instability and p53 in AID-induced c-myc-Igh translocations. *Nature* **440,** 105–109 (2006).

Ranganath, S. *et al*. Productive coupling of accessible Vβ14 segments and DJβ complexes determines the frequency of Vβ14 rearrangement. *J. Immunol.* **180**, 2339-2346 (2008).

Reina-San-Martin, B., Chen, J., Nussenzweig, A. & Nussenzweig, M. Enhanced intra-switch region recombination during immunoglobulin class switch recombination in 53BP1-/- B cells. *Eur. J. Immunol.* **37,** 235–239 (2007).

Reina-San-Martin, B., Nussenzweig, M., Nussenzweig, A. & Difilippantonio, S. Genomic instability, endoreduplication, and diminished Ig class-switch recombination in B cells lacking Nbs1. *Proc. Natl. Acad. Sci. U.S.A.* **102,** 1590–1595 (2005).

Robbiani, D. & Nussenzweig, M. Chromosome translocation, B cell lymphoma, and activation-induced cytidine deaminase. *Annu. Rev. Pathol.* **8,** 79–103 (2013).

Robbiani, D. *et al.* AID is required for the chromosomal breaks in c-myc that lead to c-myc/IgH translocations. *Cell* **135,** 1028–1038 (2008).

Robbiani, D. *et al.* AID produces DNA double-strand breaks in non-Ig genes and mature B cell lymphomas with reciprocal chromosome translocations. *Mol. Cell* **36,** 631–641 (2009).

Roddam, P. *et al.* Genetic variants of NHEJ DNA ligase IV can affect the risk of developing multiple myeloma, a tumour characterised by aberrant class switch recombination. *J. Med. Genet.* **39,** 900–905 (2002).

Roix, J., McQueen, P., Munson, P., Parada, L. & Misteli, T. Spatial proximity of translocation-prone gene loci in human lymphomas. *Nat. Genet.* **34,** 287–291 (2003).

Roldán, E. *et al.* Locus “decontraction” and centromeric recruitment contribute to allelic exclusion of the immunoglobulin heavy-chain gene. *Nat. Immunol.* **6,** 31–41 (2005).

Rooney, S. *et al.* Defective DNA repair and increased genomic instability in Artemis-deficient murine cells. *J. Exp. Med.* **197,** 553–565 (2003).

Rooney, S. *et al.* Leaky Scid phenotype associated with defective V(D)J coding end processing in Artemis-deficient mice. *Mol. Cell* **10,** 1379–1390 (2002).

Roos, W. & Kaina, B. DNA damage-induced cell death: from specific DNA lesions to the DNA damage response and apoptosis. *Cancer Lett.* **332,** 237–248 (2013).

Russo, G. *et al.* Molecular analysis of a t(14;14) translocation in leukemic T-cells of an ataxia telangiectasia patient. *Proc. Natl. Acad. Sci. U.S.A.* **86,** 602–606 (1989).

Saintigny, Y., Delacôte, F., Boucher, D., Averbeck, D. & Lopez, B. XRCC4 in G1 suppresses homologous recombination in S/G2, in G1 checkpoint-defective cells. *Oncogene* **26,** 2769–2780 (2007).

Saito, H., Hayday, A., Wiman, K., Hayward, W. & Tonegawa, S. Activation of the c-myc gene by translocation: a model for translational control. *Proc. Natl. Acad. Sci. U.S.A.* **80**, 7476-7480 (1983).

Sandlund, J. *et al.* A subtle t(3;8) results in plausible juxtaposition of MYC and BCL6 in a child with Burkitt lymphoma/leukemia and ataxia-telangiectasia. *Cancer Genet. Cytogenet.* **168,** 69–72 (2006).

Sasaki, Y. *et al*. Canonical NF-kappaB activity, dispensable for B cell development, replaces BAFF-receptor signals and promotes B cell proliferation upon activation. *Immunity* **24**, 729-739 (2006).

Schatz, D. & Swanson, P. V(D)J recombination: mechanisms of initiation. *Annu. Rev. Genet.* **45,** 167–202 (2011).

Schatz, D., Oettinger, M. & Baltimore, D. Pillars article: the V(D)J recombination activating gene, RAG-1. 1989. *J. Immunol.* **180,** 5–18 (2008).

Seifert, M., Scholtysik, R. & Küppers, R. Origin and pathogenesis of B cell lymphomas. *Methods Mol. Biol.* **971,** 1–25 (2013).

Setoodeh, R. *et al.* Double-hit mantle cell lymphoma with MYC gene rearrangement or amplification: a report of four cases and review of the literature. *Int. J. Clin. Exp. Pathol.* **6,** 155–167 (2013).

Shih, H. & Krangel, M. Chromatin architecture, CCCTC-binding factor, and V(D)J recombination: managing long-distance relationships at antigen receptor loci. *J. Immunol.* **190,** 4915–4921 (2013).

Soulas-Sprauel, P. *et al.* V(D)J and immunoglobulin class switch recombinations: a paradigm to study the regulation of DNA end-joining. *Oncogene* **26,** 7780–7791 (2007).

Spicuglia, S., Pekowska, A., Zacarias-Cabeza, J. & Ferrier, P. Epigenetic control of Tcrb gene rearrangement. *Semin. Immunol.* **22,** 330–336 (2010).

Stankovic, T. *et al.* ATM mutations in sporadic lymphoid tumours. *Leuk. Lymphoma* **43,** 1563–1571 (2002).

Stanton, L., Watt, R. & Marcu, K. Translocation, breakage and truncated transcripts of c-myc oncogene in murine plasmacytomas. *Nature* **303,** 401–406 (1983).

Stewart, G. *et al.* The DNA double-strand break repair gene hMRE11 is mutated in individuals with an ataxia-telangiectasia-like disorder. *Cell* **99,** 577–587 (1999).

Strasser, A., Harris, A., Corcoran, L. & Cory, S. Bcl-2 expression promotes B- but not T-lymphoid development in scid mice. *Nature* **368,** 457–460 (1994).

Subrahmanyam, R. & Sen, R. Epigenetic features that regulate IgH locus recombination and expression. *Curr. Top. Microbiol. Immunol.* **356,** 39–63 (2012).

Subrahmanyam, R. & Sen, R. RAGs' eye view of the immunoglobulin heavy chain gene locus. *Semin. Immunol.* **22,** 337–345 (2010).

Taccioli, G. *et al.* Ku80: product of the XRCC5 gene and its role in DNA repair and V(D)J recombination. *Science* **265,** 1442–1445 (1994).

Taylor, A., Metcalfe, J., Thick, J. & Mak, Y. Leukemia and lymphoma in ataxia telangiectasia. *Blood* **87,** 423–438 (1996).

Teng, G. *et al.* MicroRNA-155 is a negative regulator of activation-induced cytidine deaminase. *Immunity* **28,** 621–629 (2008).

Tillman, R. E., Wooley, A. L., Hughes, M., Khor, B. & Sleckman, B. Regulation of T-cell receptor beta-chain gene assembly by recombination signals: the beyond 12/23 restriction. *Immunol. Rev.* **200,** 36–43 (2004).

Truffinet, V. *et al.* The 3' IgH locus control region is sufficient to deregulate a c-myc transgene and promote mature B cell malignancies with a predominant Burkitt-like phenotype. *J. Immunol.* **179,** 6033–6042 (2007).

Tsai, A. & Lieber, M. Mechanisms of chromosomal rearrangement in the human genome. *BMC Genomics* **11 Suppl 1,** S1 (2010).

Tsai, A. *et al.* Human chromosomal translocations at CpG sites and a theoretical basis for their lineage and stage specificity. *Cell* **135,** 1130–1142 (2008).

Tsai, C., Kim, S. & Chu, G. Cernunnos/XLF promotes the ligation of mismatched and noncohesive DNA ends. *Proc. Natl. Acad. Sci. U.S.A.* **104,** 7851–7856 (2007).

Vacchio, M., Olaru, A., Livak, F. & Hodes, R. ATM deficiency impairs thymocyte maturation because of defective resolution of T cell receptor alpha locus coding end breaks. *Proc. Natl. Acad. Sci. U.S.A.* **104,** 6323–6328 (2007).

Vanasse, G. *et al.* Genetic pathway to recurrent chromosome translocations in murine lymphoma involves V(D)J recombinase. *J. Clin. Invest.* **103,** 1669–1675 (1999).

Vettermann, C. & Schlissel, M. Allelic exclusion of immunoglobulin genes: models and mechanisms. *Immunol. Rev.* **237,** 22–42 (2010).

Vitolo, U. *et al.* T acute lymphoblastic leukemia in ataxia-telangiectasia. Report of a case characterized by monoclonal antibodies. *Haematologica* **69,** 695–700 (1984).

Vousden, K. & Prives, C. Blinded by the Light: The Growing Complexity of p53. *Cell* **137,** 413–431 (2009).

Vuong, B. & Chaudhuri, J. Combinatorial mechanisms regulating AID-dependent DNA deamination: interacting proteins and post-translational modifications. *Semin. Immunol.* **24,** 264–272 (2012).

Wake, N., Minowada, J., Park, B. & Sandberg, A. A. Chromosomes and causation of human cancer and leukemia. XLVIII. T-cell acute leukemia in ataxia telangiectasia. *Cancer Genet. Cytogenet.* **6,** 345–347 (1982).

Wang, J. & Boxer, L. Regulatory elements in the immunoglobulin heavy chain gene 3'-enhancers induce c-myc deregulation and lymphomagenesis in murine B cells. *J. Biol. Chem.* **280,** 12766–12773 (2005).

Wang, J. *et al.* Mechanisms promoting translocations in editing and switching peripheral B cells. *Nature* **460,** 231–236 (2009).

Wang, J. *et al.* Oncogenic transformation in the absence of Xrcc4 targets peripheral B cells that have undergone editing and switching. *J. Exp. Med.* **205,** 3079–3090 (2008).

Ward, I. *et al.* 53BP1 cooperates with p53 and functions as a haploinsufficient tumor suppressor in mice. *Mol. Cell. Biol.* **25,** 10079–10086 (2005).

Ward, I. *et al.* 53BP1 is required for class switch recombination. *J. Cell Biol.* **165,** 459–464 (2004).

Wilda, M. *et al.* Level of MYC overexpression in pediatric Burkitt's lymphoma is strongly dependent on genomic breakpoint location within the MYC locus. *Genes. Chromosomes Cancer* **41,** 178–182 (2004).

Williams, R. *et al.* Amplification of complex gene libraries by emulsion PCR. *Nat. Methods* **3,** 545–550 (2006).

Wu, C. *et al.* Dramatically increased rearrangement and peripheral representation of Vbeta14 driven by the 3'Dbeta1 recombination signal sequence. *Immunity* **18,** 75–85 (2003).

Xu, Y. & Baltimore, D. Dual roles of ATM in the cellular response to radiation and in cell growth control. *Genes Dev.* **10,** 2401-2410 (1996).

Yan, C. *et al.* IgH class switching and translocations use a robust non-classical end-joining pathway. *Nature* **449,** 478–482 (2007).

Yan, J. & Jetten, A. RAP80 and RNF8, key players in the recruitment of repair proteins to DNA damage sites. *Cancer Lett.* **271,** 179–190 (2008).

Yancopoulos, G. & Alt, F. Developmentally controlled and tissue-specific expression of unrearranged VH gene segments. *Cell* **40,** 271–281 (1985).

Zarrin, A. *et al.* Antibody class switching mediated by yeast endonuclease-generated DNA breaks. *Science* **315,** 377–381 (2007).

Zha, S. *et al.* ATM-deficient thymic lymphoma is associated with aberrant tcrd rearrangement and gene amplification. *J. Exp. Med.* **207,** 1369–1380 (2010).

Zha, S., Alt, F., Cheng, H., Brush, J. & Li, G. Defective DNA repair and increased genomic instability in Cernunnos-XLF-deficient murine ES cells. *Proc. Natl. Acad. Sci. U.S.A.* **104,** 4518–4523 (2007).

Zhang, M. & Swanson, P. V(D)J recombinase binding and cleavage of cryptic recombination signal sequences identified from lymphoid malignancies. *J. Biol. Chem.* **283,** 6717–6727 (2008).

Zhang, Y. & Rowley, J. D. Chronic myeloid leukemia: current perspectives. *Clin. Lab. Med.* **31,** 687–698, x (2011).

Zhang, Y. *et al.* Spatial organization of the mouse genome and its role in recurrent chromosomal translocations. *Cell* **148,** 908–921 (2012).

Zhang, Y. *et al.* The role of mechanistic factors in promoting chromosomal translocations found in lymphoid and other cancers. *Adv. Immunol.* **106,** 93–133 (2010).

Zhao, J., Bacolla, A., Wang, G. & Vasquez, K. Non-B DNA structure-induced genetic instability and evolution. *Cell. Mol. Life Sci.* **67,** 43–62 (2010).

Zhu, C. *et al.* Unrepaired DNA breaks in p53-deficient cells lead to oncogenic gene amplification subsequent to translocations. *Cell* **109,** 811–821 (2002).

Zoete, V., Irving, M., Ferber, M., Cuendet, M. & Michielin, O. Structure-based, rational design of T cell receptors. *Front. Immunol.* **4,** 268 (2013).